Identification of Deubiquitylases Involved in the Regulation of Adherens Junction Components

Jia Lih Wong

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Philosophy

October 2012

Identification of Deubiquitylases Involved in the Regulation of Adherens Junction Components

Jia Lih Wong

Abstract

Protein ubiquitylation represents a versatile mechanism to regulate multiple cellular processes, including protein turnover, transcription, cell signaling, membrane trafficking and DNA damage repair. The attachment of ubiquitin moieties to a target substrate is sequentially catalysed by three enzymes, namely E1, E2 and E3, and this process can be reversed by the deubiquitylases (DUBs). This research project seeks to decipher the involvement of DUBs in regulation of the components of the adherens junction (AJ), E-cadherin and β-catenin, in particular, which are the core components of the AJ.

It is found in MCF7 cells, a breast cancer cell line, that E-cadherin undergoes constitutive turnover via the lysosomal pathway, as indicated by the accumulation of an 80kDa fragment of E-cadherin degradative fragment following treatment of MCF7 cells with Folimycin, a lysosomal v-ATPase inhibitor. Using the ratio of the full length and 80kDa degradative fragment as a biochemical readout for the trafficking status of E-cadherin, a siRNA human DUB library screen was performed and USP38 was identified as a leading candidate. siRNA depletion of USP38 led to increase in the full length to 80kDa E-cadherin ratio. Notably, the loss of USP38 led to significant loss of total E-cadherin level (sum of full length E-cadherin and 80kDa E-cadherin fragment).

The loss of E-cadherin at the protein level following USP38 knockdown was not accompanied by a decrease in its mRNA level or increase in degradation rate, suggesting that translational effects could operate. Preliminary characterisation of this enzyme showed that it has both nuclear and cytosolic localisation. Mapping of nuclear localisation determinants revealed a role for a 46 amino acid insertion within the core catalytic domain, and the last 100 amino acid on C-terminus of USP38.

In a separate siRNA DUB library screen performed on A549 cells, the established tumour suppressor BAP1 was identified as a DUB regulating β-catenin. siRNA depletion of BAP1 resulted in a loss of β-catenin and loss of cell-cell contacts, whereas its overexpression led to an increase in β-catenin level, indicating a positive regulatory role of BAP1 on β-catenin level.

To facilitate further screening efforts and to provide an extra layer of validation to my own screens, a human DUB endonuclease-prepared small interfering RNA (esiRNA) library was developed. EsiRNA is a complex mixture of siRNAs generated by random cleavage of long double-stranded RNA and is argued to be able to circumvent the problem of siRNA-induced off-target effects. EsiRNAs were made for 91 out of the currently predicted 93 human DUBs. Quality control experiments were carried out upon a sample of selected DUBs and optimisation of experimental conditions for efficient knockdown in various cell lines was explored.

Table of Contents

List of Figures

List of Tables

Appendices

Abbreviations

AA	Amino acid
ABA	Aureobasidin
AJ	Adherens Junctions
AMSH	Associated molecule with SH3 domain of STAM
AMSH-LP	AMSH-like protein
Atg12	Autophagy related protein 12
ATP	Adenosine triphosphate
APC	Adenomatous Polyposis Coli
APF1	ATP-dependent Proteolysis Factor 1
bHLH	Basic Helix-loop-helix
CBP	CREB-binding protein
ChIP	Chromatin immunoprecipitation
CK1	Casein Kinase 1
CRL	Cullin RING ligases
CUE	Coupling of Ubiquitin conjugation to ER degradation domain
DAPI	4’,6-diamidino-2-phenylindole
DNA	Deoxyribonucleic acid
DMEM	Dulbecco’s Modified Eagle Medium
DO	Different Output
dsRNA	Double Stranded-RNA
Dvl	Dishevelled
EEA1	Endosome Antigen 1
EGFR	Epidermal Growth Factor Receptor
EMT	Epithelial-to-Mesenchymal Transition
ESCRT	Endosomal Sorting Complex Required for Transport
ER	Endoplasmic reticulum
ERAD	Endoplasmic reticulum associated degradation
esiRNA	Endonuclease-prepared siRNA
FAP	Familial Adenomatous Polyposis

FBS	Foetal Bovine Serum
GO	Gene Ontology
GST	Glutathione S Transferase
HCF1	Host Cell Factor 1
HECT	Homologous to E6AP Carboxy Terminus
HGF	Hepatocyte Growth Factor
ISG15	Interferon stimulated gene 15
IVT	In Vitro Transcription
JAMMs	JAB1/MPN/Mov4 metalloenzymes
MDCK	Madin Darby Canine Kidney
MJDs	Machado Joseph Disease proteins
MMP	Matrix Metalloproteases
MoaD	Molybdopterin converting factor subunit 1
MoeB	Molybdopterin Biosynthetic Enzyme B
MPED	Membrane Proximal Extracellular Domain
mRNA	Messenger ribonucleic acid
MVB	Multivesicular Body
Nedd8	Neural precursor cell expressed, developmentally down-regulated 8
NMR	Nuclear Magnetic Resonance
ORF	Open Reading Frame
OTUD	OTU-domain
OTUs	Ovarian Tumour Proteases
PCP	Planar Cell Polarity
PCR	Polymerase Chain Reaction
RB	Retinoblastoma Protein
RING	Really Interesting New Gene
RNAi	RNA Inteference
RTK	Receptor Tyrosine Kinases
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide
SILAC	Stable isotope labeling by amino acid in cell culture

siRNA	Small interfering ribonucleic acid
S.O.C.	Super optimal broth catabolite repression
STAM	Signal Transducing Adaptor Molecule
SUMO	Small ubiquitin-like modifier
TCF/LEF	T-cell factor/Lymphoid enhancer factor
TGN	Trans Golgi Network
ThiS	Thiamine Biosynthesis Protein S
TKB	Tyrosine Kinase Binding
TNFα	Tumour Necrosis Factor Alpha
TRAIL	TNF-related apoptotis inducing ligand
TW	Terminal Web
Ub	Ubiquitin
UBA	Ubiquitin Associated Domain
Ub-AMC	Ubiquitin COOH-terminal 7-amino-4-methylcoumarin
UBC	Ubiquitin-Conjugating
UBD	Ubiquitin Binding Domains
UBZ	Ubiquitin-Binding Zinc finger
UCHs	Ubiquitin C-terminal Hydrolases
UFD	Ubiquitin-Fold Domain
UIM	Ubiquitin Interacting Motif
UPS	Ubiquitin-Proteasome System
USPs	Ubiquitin Specific Proteases
UV	Uveal Melanoma
UTR	Untranslated Region
VHL	Von Hippel Lindau
YY1	Yin Yang 1

Acknowledgement

“结束只是另一个开始……. ”

“Ending is beginning…….“

It was an incredible four years that I owe many an enormous debt of gratitude.

First and foremost, an immense thank you to my supervisors, Michael Clague and Sylvie Urbé, for giving me the most precious learning opportunity that I could ever ask for. Thanks very much too for being a constant source of knowledge, inspiration and motivation. It was an overwhelmingly rewarding and invaluable intellectual experience.

Thanks to Judy Coulson and Ian Prior for your guidance and inputs for my research project. Thanks to Ronald Brian Beechey for offering lots of scientific food for thoughts, which embody the precious legacies of the “old-school” scientists, and not least, for being a lovely friend. Thanks to my thesis supervisor Alan Morgan for being encouraging and for your advices. My thanks also go to past and present members of the lab: Carolyn Levene, Cormac O’Reilly, Sebastian Hayes, Christopher Thorne, Monika Chojnowska, Sara Cadeco, Rebecca Eccles, Ewan MacDonald, Craig Mageean, Anna Newlaczyl, Dead Hammond, Alison Beckett, Simon Oliver, Veronica Aran, Joseph Sacco, Viktor Malec and Maria Hernandez-Valladares, for work-related assistance and for making the lab a pleasant working environment. Particular thanks to Amos Liang and Yvonne Tang for helping me in various ways in the lab and for being lovely and supportive flatmates. Many thanks to Claire Heride for helping me with Y2H experiments. Many thanks to Monica Faronato for being always responsive to requests for helps. Many thanks to Han Liu for being a source of knowledge and ideas. Special thanks to Jasminka Omerovic for all the scientific discussions, for the friendly conversations and for being a constant source of support at all times.

I am extremely grateful to Cancer Research UK for granting me the studentship and enabling me to carry out my research work in Liverpool.

Thanks very much too to all the teachers and mentors that I have met in my life, for I am who I am because of you.

I would also like to express my utmost gratitude to all the friends in Liverpool and around the UK, who keep me in company through ups and downs, and fill my life with laughters and inspirations. Thanks to Rui Ying, Tong Tong, Kwang Heng, Vicky, Liu Yan, Ya Wen and Lih Tyng for staying with me over the last 7 years in the UK. In particular, I would like to thank Chloe Tan and Yih Jian Liew for their supports and helps when they were very much needed. Many thanks to my fellow 24 Festival Drum teammates, who spiced up the very last year of my PhD. Also, special thanks to Yi Cheng and Yuan Yuan for the joys and comforts you bring me as always.

Last but not least, I would like to dedicate this thesis to my family. Thank you very much Mum and Dad, Jia Yueh and Jia Yih for all the loves and supports you provided me all the way long. I LOVE YOU!

Chapter 1

Introduction

1.1 Adherens Junctions

The integrity of tissue architecture and functional diversification of a multilayered metazoan organism is dependent on the physical coupling of neighbouring cells as well as interaction between the cells and their environment (Franke, 2009; Yap *et al.*, 1997). This is achieved by four major types of cell-cell junctions, namely adherens junctions (AJ), desmosomes, tight junctions and gap junctions (Franke, 2009). Among these, the AJ is the key determinant of cell-cell adhesion and it provides physical links between membrane proteins and cytoskeletal components at discrete contact regions between cells (Niessen and Gottardi, 2008).

By performing ultrastructural analysis of animal tissues using electron microscopy, Farquhar and Palade provided the first detailed description of zonular AJ, which, together with desmosomes and tight junctions, forms a tripartite junctional complex between the apices of a variety of simple epithelia (Farquhar and Palade, 1963). AJ, as described by Farquhar and Palade, comprises of apposed membrane separated by an intercellular space of ~200Å with dense plaques of cytoplasmic microfilaments on both sides. A subsequent study, using deep-etch electron microscopy, identified actin filaments as major components of the plaques and revealed numerous cylindrical projections from the apposing membrane, filling the intercellular space at the AJ (Hirokawa and Heuser, 1981).

Based on these ultrastructural observations, one would anticipate there are at least three classes of proteins present at the AJ: (i) adhesive protein units which span the intercellular space, thereby providing physical link between neighbouring

cells; (ii) a cytoskeletal network onto which the adhesive proteins are anchored, and (iii) proteins which mediate association of the adhesive proteins with the cytoskeletal network (Niessen and Gottardi, 2008).

1.1.1 Functions of AJ

At the cellular and molecular level, AJ is important for processes including initiation and stabilisation of cell-cell adhesion, modulation of the actin cytoskeleton, intracellular signalling events and regulation of transcriptional events (Niessen, 2007; Hartsock and Nelson, 2008).

Evidence suggests that E-cadherin-based AJ is a prerequisite for the assembly of the other intercellular junctions (Gumbiner *et al.*, 1988). The conditional inactivation of E-cadherin gene in mice resulted in mislocalisation of tight junctional proteins and hence a leaky epidermal water barrier that led to perinatal death (Tunggal *et al.*, 2005). Multiple studies have also demonstrated that ablation of cadherin-based AJ disrupts or delays the assembly of desmosomes (Lewis *et al.*, 1994; Lewis *et al.*, 1997).

1.1.3 Core Components of AJ – the Cadherin/Catenin Complex

There are two alternative adhesive protein units of AJ, namely cadherin/catenin complex and nectin/afadin complex (Niessen and Gottardi, 2008). Cadherin, specifically, E-cadherin (due to its epithelial origin) was the first adhesive component of AJ that was discovered (Takeichi 1977; Yoshida-Noro *et al.*, 1984). It was shown to be essential for morula compaction during early stage of embryogenesis, a cell-cell adhesion dependent event, and that this process was Ca^{2+} dependent, hence the name cadherin. Since it is the first member of the cadherin superfamily to be identified, E-cadherin, together with its subfamily members, P-

cadherin (placenta), N-cadherin (neuronal), VE-cadherin (vasoendothelial), are classified as classic cadherins (Angst *et al.*, 2001; Pettitt 2005).

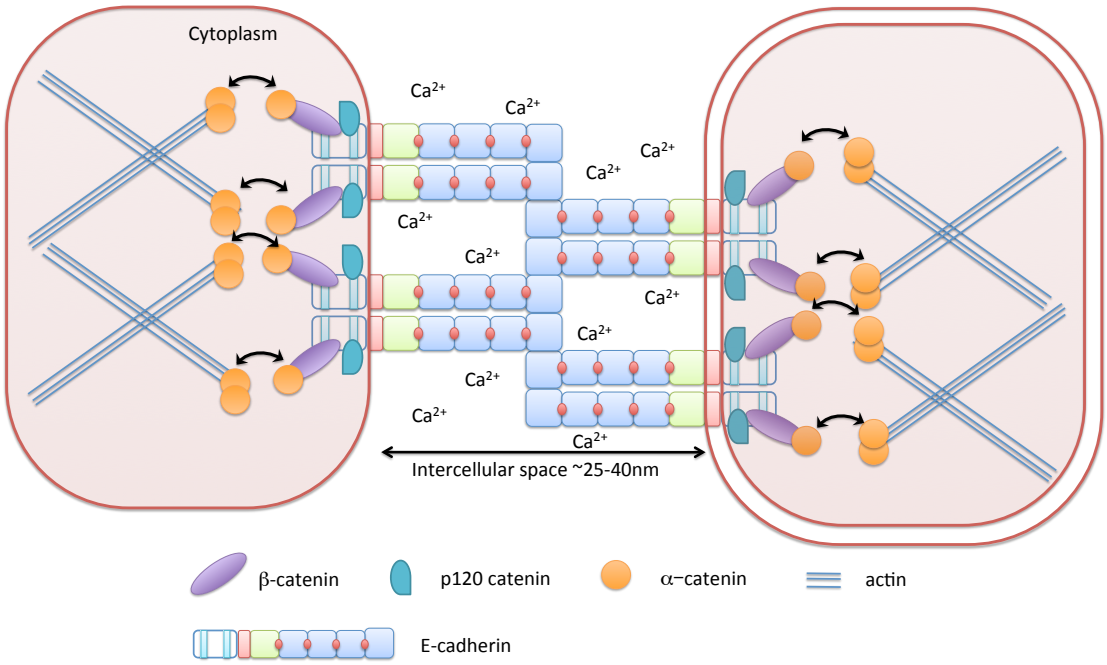

Figure 1.1. Core components of adherens junction complex. E-cadherin molecules expressed on neighbouring cells form homophilic interaction in a calcium-dependent manner, forming the core of the AJ. On the cytoplasmic tail, p120-catenin binds to the juxtamembrane domain, stabilising the AJ on the plasma membrane. β-catenin binds to the cytoplasmic tail of E-cadherin, and interacts with monomeric α-catenin. With high concentration of the cadherin-catenin complex at the AJ, the α-catenin switches between monomeric and homo-dimeric form, where the dimeric α-catenin helps anchor the AJ complex to actin cytoskeleton.

The cytoplasmic domain of cadherin is associated with three armadillo repeat proteins, namely β-catenin, α-catenin and p120-catenin (Figure 1.1) (Ozawa *et al.*, 1989). This association is important for the adhesion and signalling activities mediated by cadherin (Niessen and Gottardi, 2008; van Roy and Berx, 2008). The binding of β-catenin to cadherin is a prerequisite for cell adhesion, as the association of the two proteins at endoplasmic reticulum (ER) facilitates the transport of the complex to cell surface (Chen *et al.*, 1999) and the binding of β-catenin masks a PEST sequence (a proteasomal degradation motif) on the cytoplasmic tail of E-cadherin,

which, if exposed, renders the molecule for proteolytic processing (Huber *et al.*, 2001).

α-catenin has long been thought to provide the physical link between the cadherin/β-catenin complex and the actin cytoskeleton, since it was demonstrated to bind actin filaments (Rimm *et al.*, 1995) and β-catenin in vitro (Aberle *et al.*, 1994). However, Yamada and colleagues demonstrated that α-catenin cannot simultaneously bind β-catenin and actin (Yamada *et al.*, 2005). This is because α-catenin switches between a monomeric and a homo-dimeric state, where the monomeric α-catenin binds only to β-catenin (but not actin) whereas the dimeric α-catenin binds only to actin (but not β-catenin) (Figure 1.1); and that the region required for binding to β-catenin overlaps with the region for α-catenin homodimerisation (Pokutta and Weis, 2000; Yamada *et al.*, 2005). Moreover, *in vitro* studies show that a 10-fold higher concentration of monomeric α-catenin is required for its dimerisation (Drees *et al.*, 2005). Putting this into the context of AJ assembly, it therefore indicates requirement for local concentration of monomeric α-catenin, possibly by clustering of α-catenin-bound cadherin/β-catenin complex at the plasma membrane, to allow subsequent dimerisation of α-catenin to mediate actin-filament organisation (from branched to bundled) at cell-cell contact site. Such a working model allows dynamic coupling of the adherens junction complex assembly and actin remodelling (Niessen and Gottardi, 2008).

P120-catenin binds to a highly conserved 8-amino-acid motif (YDEEGGGE) (Ferber *et al.*, 2002) within the juxtamembrane domain (JMD) of E-cadherin (Reynolds *et al.*, 1996; Ohkubo and Ozawa, 1999). Peifer and Yap proposed the role of p120-catenin as a gatekeeper of classical cadherins, which dictates the stability of the cadherins on plasma membrane and fate of cadherin during endocytic trafficking (Peifer and Yap, 2003; Xiao *et al.*, 2007). In a colon cancer cell line SW48, which harbors a mutation leading to the truncation of p120-catenin, there is a low level of E-cadherin protein despite high level of E-cadherin messenger ribonucleic acid (mRNA) (Ireton *et al.*, 2002). The prominent loss of E-cadherin could be reverted by reintroduction of wildtype p120-catenin, suggesting regulation of E-cadherin protein

turnover by p120-catenin. In line with this observation, the small siRNA (small interfering ribonucleic acid) depletion of p120-catenin in cells expressing wildtype p120-catenin results in loss of classical cadherins, including E-cadherin, N-cadherin and VE-cadherin (Davis *et al.*, 2003; Xiao *et al.*, 2003). The loss of p120-catenin results in internalisation of cadherin molecules on the plasma membrane and their subsequent trafficking to the lysosomes for proteolytic degradation.

1.1.2.1 E-cadherin

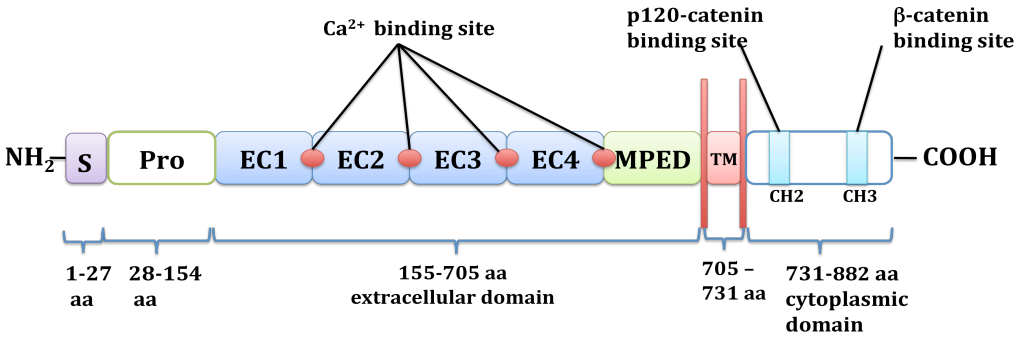

Figure 1.2. Domain architecture of E-cadherin. E-cadherin is synthesised as precursor polypeptide, with a propeptide and ER signal peptide at the N-terminus. A mature protein consists of 5 tandemly repeated domains: 4 extracellular cadherin repeats (EC1 to EC4) followed by membrane proximal extracellular domain (MPED). A transmembrane domain allows the protein to be embedded across the membrane. The cytoplasmic domain of E-cadherin contains binding sites for p120-catenin and β-catenin, which are named CH2 and CH3 respectively.

E-cadherin is synthesised as a precursor polypeptide, with a propeptide of about 154 amino acids (aa) on its N-terminus (Figure 1.2) (van Roy and Berx, 2008). The propeptide has an ER signal peptide for import into the ER and the 127 aa C-terminal to the signal peptide is cleaved off as the polypeptide matures, giving rise to a final mature polypeptide of 728 aa. Mature E-cadherin is a single-span transmembrane glycoprotein.

The 550 aa extracellular domain consists of 5 tandemly repeated domains, where the first 4 domains are known as extracellular cadherin repeats (EC1 to EC4).

The EC repeat is characteristic of all the cadherins but only varies with number among the different members of cadherin superfamily. The fifth domain is known as the membrane proximal extracellular domain (MPED). This domain is characterised by 4 highly conserved cysteine residues, which form disulphide bridges that are important for the adhesive properties of the molecule (Ozawa *et al.*, 1990). There are Ca^{2+} binding sites in between the EC repeats and the binding of Ca^{2+} helps ensure the rigidity of the molecule in an elongated conformation (Yap *et al.*, 1997). The extracellular domain of E-cadherin is responsible for homophilic interaction of E-cadherin molecules expressed on neighbouring cells (Nose *et al.*, 1988) and the specificity of interaction is determined by a highly conserved HAV motif and the 2 amino acids flanking the motif in EC1 (Blaschuk *et al.*, 1990; Tsuji *et al.*, 1990).

The cytoplasmic domain of E-cadherin also plays an equally important role in determining cell adhesion through its interaction with the catenins (Ozawa *et al.*, 1989). The juxtamembrane domain, also known as cadherin homology domain 2 (CH2) of 30-35 aa, contains a sequence motif for binding of p120-catenin, whereas the CH3 domain allows binding of β-catenin (Rimm and Morrow 1994). Within the CH3 domain, there is a cluster of 8 serine residues between amino acids 684 to 699 (corresponding to amino acids 838 – 853 in preprotein of E-cadherin), which is important for regulation of interaction between E-cadherin and β-catenin (Lickert *et al.*, 2000). The phosphorylation of 3 serines residues, namely Ser684, Ser686 and Ser692 by Casein Kinase II (CK II) and glycogen synthase kinase 3 β (GSK3β) significantly strengthens interaction between the two proteins.

1.1.3.1.1 Transcriptional regulation of E-cadherin

E-cadherin is encoded by *CDH1* gene. Our understanding of the transcriptional regulation of E-cadherin owes much to the elucidation of regulatory sequences of the mouse *CDH1* gene, which is characterised by modules of binding regions for different transcription factors (Figure 1.3) (van Roy and Berx, 2008; Comijn *et al.*, 2001). On the 5-prime (5') end of *CDH1* gene, there are 2 positive regulatory

elements, namely a CCAAT Box and 2 GC boxes. In mouse, E-cadherin transcription can be initiated by the transcription factor AP2 upon binding to the GC box (Hennig *et al.*, 1996). Moreover, the tumour suppressor, retinoblastoma protein (Rb) and the proto-oncogene, c-Myc can act as co-activators of AP2 in epithelial cells to induce E-cadherin expression (Batsche *et al.*, 1998; Decary *et al.*, 2002). The binding of WT1, a tumour suppressor gene that is mutated in Wilm's tumor, to GC Box of E-cadherin can induce fibroblasts to undergo differentiation and acquire epithelial phenotypes, when WT1 is ectopically expressed (Hosono *et al.*, 2000).

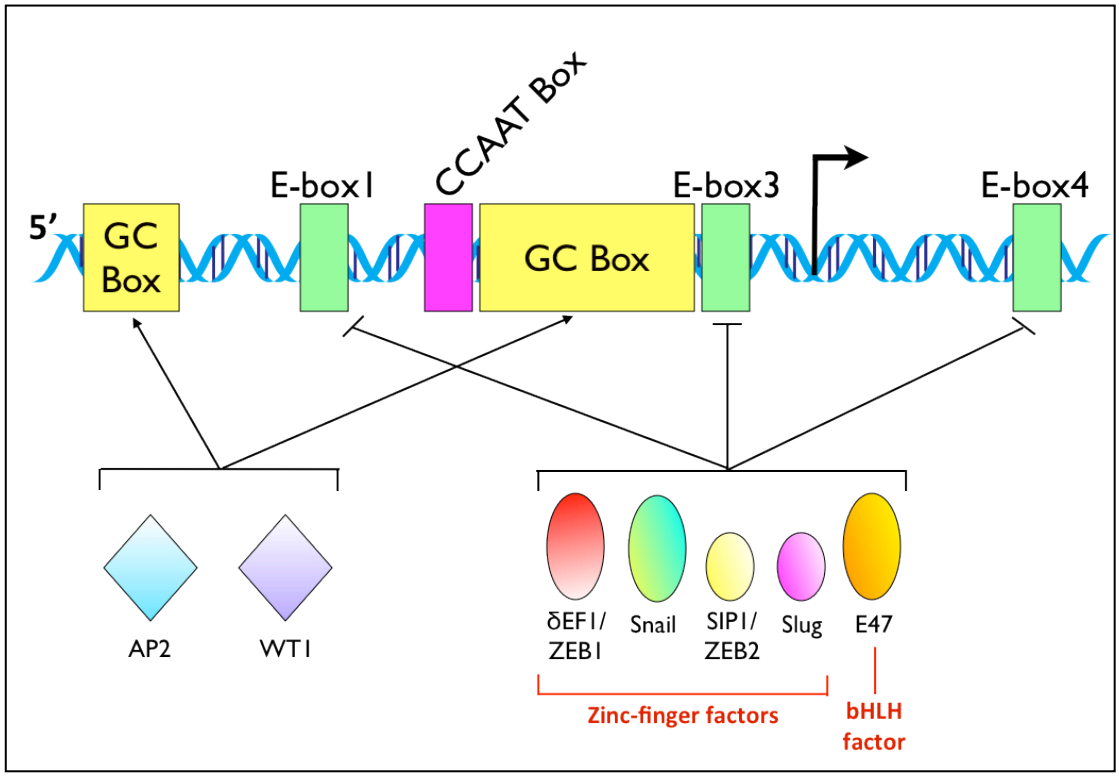

Figure 1.3. Transcriptional regulation of E-cadherin. The regulatory sequence, most of which are upstream of the 5' end, of the *CDH1* gene encoding for E-cadherin exhibits a modular structure. The CCAAT box and GC box are positive regulatory sequences that enable E-cadherin expression, upon binding of transcription factor such as AP2 and WT1 to the GC box. The E-boxes are negative regulatory sequences, the binding of which by transcriptional repressors such as δEF1/ZEB1, Snail, SIP1/ZEB2, Slug and E47 prevents E-cadherin expression.

Birchmeier group has also reported presence of a palindromic sequence (termed E-box) between position about 75 to 86 base 5' of *CDH1* gene, which

potentiates activity of E-cadherin promoter in epithelial but inhibits promoter activity in non-epithelial cells (Behrens *et al.*, 1991). Notably, the inactivation of the E-box in mesenchymal cells induces E-cadherin promoter activity, suggesting the presence of an E-cadherin transcriptional repressor in non-epithelial cells. Using one-hybrid screening with the E-box sequence as bait, Cano and colleagues identified Snail as the first potent transcriptional repressor of E-cadherin (Cano *et al.*, 2000). This was then followed by identification and characterisation of δEF1/ZEB1 (Eger *et al.*, 2005), SIP1/ZEB2 (Comijn *et al.*, 2001), Slug (Hajra *et al.*, 2002; Bolos *et al.*, 2003) and E12/E47 (Perez-Moreno *et al.*, 2001), which exert their transcriptional repression activity via the E-box.

1.1.3.1.2 Endocytic regulation of E-cadherin

Throughout the life-span of an E-cadherin molecule, a significant amount of this time is attributed to trafficking via the different subcellular membrane compartments (Figure 1.4).

Upon its biosynthesis in the ER, E-cadherin is coupled to β-catenin and this association is crucial for ER exit of the complex and subsequent baso-lateral trafficking to plasma membrane via the trans Golgi network (TGN) (Chen *et al.*, 1999). The correct sorting of the E-cadherin-β-catenin complex to the plasma membrane is dependent on a highly conserved dileucine motif at the membrane proximal cytoplasmic domain (Miranda *et al.*, 2001; Miranda *et al.*, 2003). Deletion of this motif results in missorting of the E-cadherin-β-catenin complex and consequently the disruption of cell polarity.

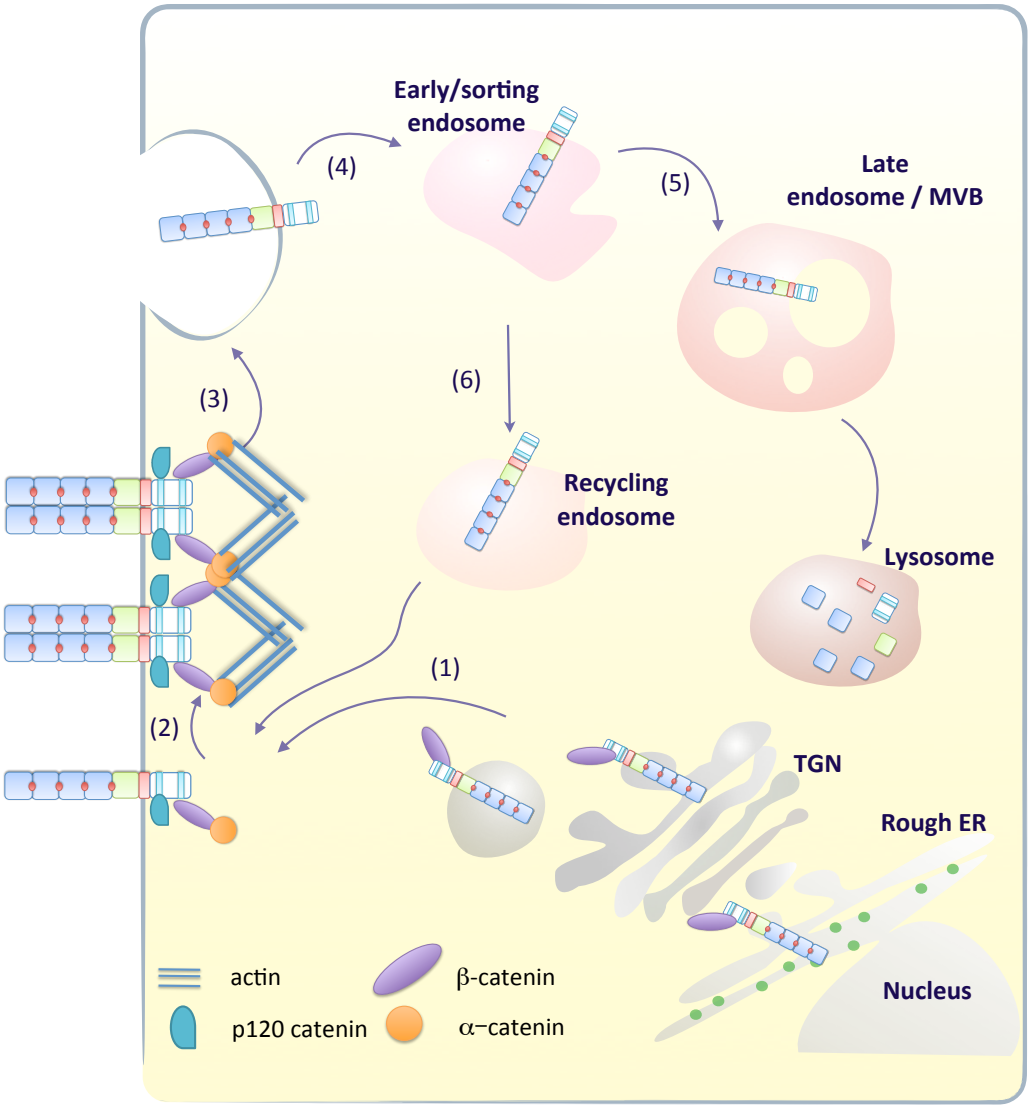

Figure 1.4. Endocytic regulation of E-cadherin. (1) Upon its synthesis at the rough endoplasmic reticulum (ER), E-cadherin is associated with β-catenin and this association enables ER exit of the complex, and subsequent trafficking to the plasma membrane via the trans-Golgi network (TGN). At the plasma membrane, the E-cadherin-β-catenin is stabilised by the binding of p120-catenin at the JMD of E-cadherin, thereby allowing assembly of the AJ. (3) The AJ is a dynamic structure that undergoes constitutive or inducible internalisation via different endocytic pathways (clathrin-dependent or –independent pathways). (4) Following endocytosis, the E-cadherin enters the early endosomes, and is later either (5) sorted into the late endosomes/multivesicular body (MVB) for subsequent degradation via the lysosomal pathway or is (6) transported to the recycling endosome for recycling back to the plasma membrane.

The multiple endocytic routes can commit E-cadherin to very different fates: recycling back to the membrane, temporary sequestration in cellular compartments

such as sorting and recycling endosomes or routing to late endosomes for subsequent lysosomal degradation (Bryant and Stow, 2004). Superimposed onto these possible routes of trafficking are the physiological events that trigger E-cadherin internalisation in the first place. It is reasonable to deduce that a particular physiological condition induces E-cadherin internalisation via a particular uptake route, which will lead to a definite fate of E-cadherin. For example, internalised E-cadherin by macropinocytosis following stimulation of non-confluent MCF7 cells is recycled back to the plasma membrane (Bryant *et al.*, 2007). While such deduction is plausible, there are reports of E-cadherin, whether internalised via the clathrin-dependent or independent pathways, entering cellular compartment positive for early endosome antigen 1 (EEA1) (Xiao *et al.*, 2003). Hence, it suggests that the fate of E-cadherin is determined at or after leaving the early endosomal compartment (Yap *et al.*, 2007). Early endosome is predominantly a tubulovesicular membrane compartment, which serves as a major platform for transiting plasma membrane proteins (Williams and Urbé, 2007). It is believed that the structure and components of this compartment undergoes progressive changes and eventually becomes late endosomes characterised by the presence of multiple intraluminal vesicles. Therefore, the late endosome is also known as multivesicular body (MVB). The MVB eventually fuses with lysosomes and its contents are therefore hydrolytically digested. Ubiquitylation is the best-characterised sorting signal, which commits membrane proteins into the degradative MVB pathway (Clague *et al.*, 2012).

Ubiquitylation of E-cadherin was first demonstrated in MDCK cells by Birchmeier's group (Fujita *et al.*, 2002). It was shown that the tyrosine phosphorylation of E-cadherin by vSrc, was accompanied by its ubiquitylation as well, and the consequence of which was the disassembly of the E-cadherin-catenin complex and endocytosis of E-cadherin. By virtue of yeast-two-hybrid technique, using an ingeniously designed chimeric protein of E-cadherin and active Met kinase to enable E-cadherin phosphorylation in yeast, Fujita and colleagues identified a c-cbl like really interesting new gene (RING) E3 ligase and named it Hakai, which means destruction in Japanese. Hakai specifically interacts with tyrosine phosphorylated E-cadherin and, together with E1 and E2 ubiquitylating enzymes,

catalyses ubiquitylation of E-cadherin (Figure 1.5). The ubiquitylated E-cadherin is then endocytosed and targeted for lysosomal degradation.

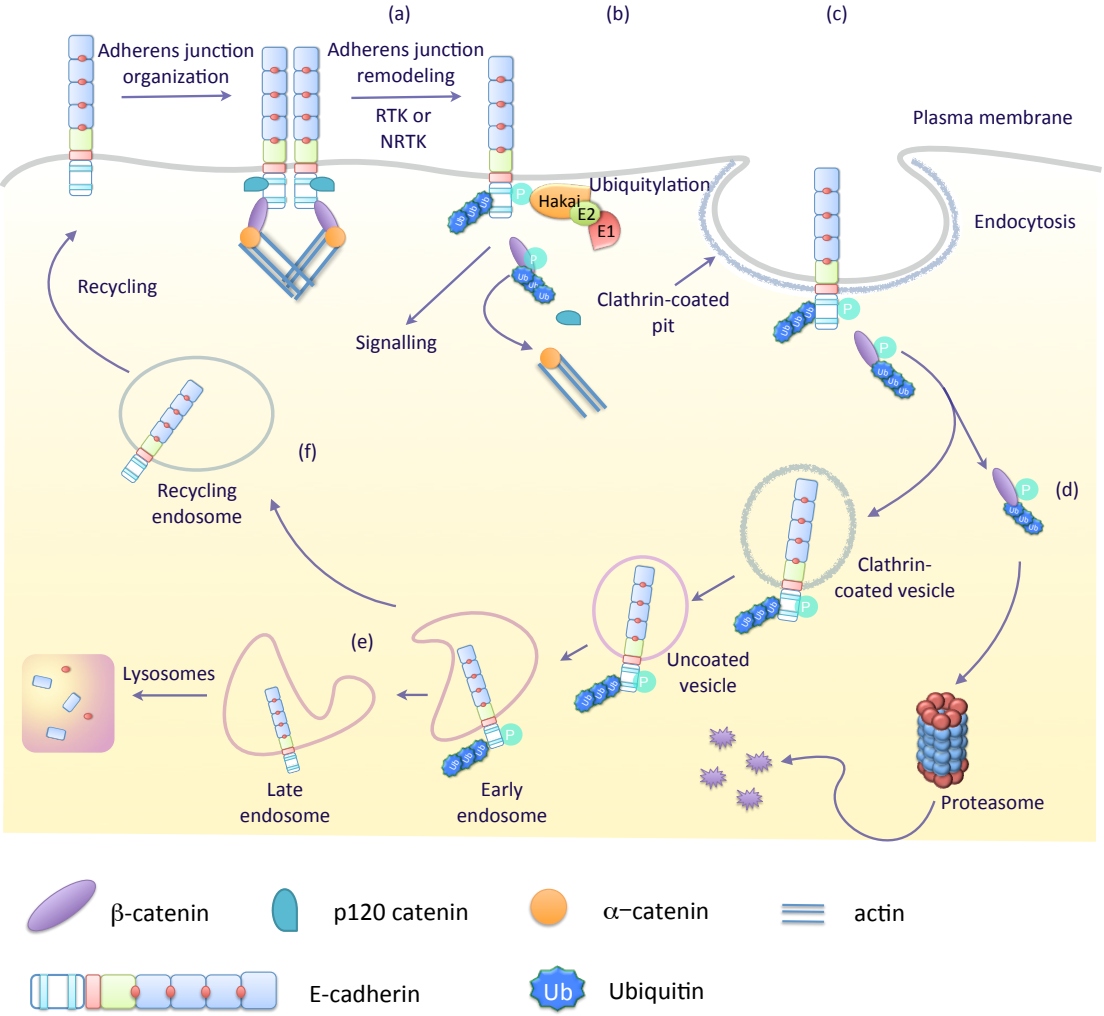

Figure 1.5. Ubiquitin-dependent regulation of E-cadherin. (a) Following stimulation by receptor tyrosine kinase (RTK) or non-receptor tyrosine kinase (NRTK), E-cadherin and β-catenin are phosphorylated, leading to dissolution of the AJ. (b) The phosphorylated E-cadherin is then recognised by Hakai, the E3 ligase, which catalyses its ubiquitylation with the presence of E1 and E2 enzymes. (c) The polyubiquitylated E-cadherin is subsequently internalised via the clathrin dependent pathway, (d) while the polyubiquitylated β-catenin is targeted for proteasomal degradation. The internalised E-cadherin subsequently localises to the early endosome, where its fate was decided between (e) sorting to the late endosome for lysosomal degradation or (f) recycling back to the plasma membrane for AJ assembly.

Another indication of ubiquitylation-dependent regulation of E-cadherin involves the DUB Fat facet in mouse (FAM), also known as USP9X. USP9X was first found to be a developmentally regulated gene (Taya *et al.*, 1998; Taya *et al.*, 1999)

and was identified to interact with and stabilise β-catenin (Taya *et al.*, 1999). Subsequent study by Murray *et al.* (2004) showed that USP9X colocalised with E-cadherin and β-catenin at multiple points of protein trafficking in subconfluent epithelial cells, but did not colocalise with E-cadherin and β-catenin at plasma membrane upon formation of stable AJ in a confluent layer of epithelial cells. Such association was confirmed biochemically too, where USP9X was found to form a complex with E-cadherin and β-catenin in a Triton X-100 soluble fraction (E-cadherin-β-catenin complex not involved in AJ formation is Triton X-100 soluble) and the DUB was not found in a Triton X-100 insoluble fraction. Together these results suggested that USP9X is involved in the trafficking of the E-cadherin-β-catenin complex in their biosynthetic pathway but is not involved in their regulation upon AJ formation. Of note, this represents a more general role of USP9X in protein trafficking as it colocalises on endocytic vesicles with multiple other proteins, such as alpha actinin and myosin II, rather than a specific regulatory relationship between USP9X and the E-cadherin-β-catenin complex.

1.1.3.1.3 Mechanisms of dysregulation of E-cadherin expression and cancer

Given the pivotal roles of E-cadherin in tissue morphogenesis and development, it is no surprise that the loss or dysregulation of E-cadherin expression has severe pathological consequences, such as that in cancer. Birchmeier's group provided one of the first experimental evidence that induction of E-cadherin loss confers epithelial cells with invasive properties (Behrens *et al.*, 1989). Subsequently, there were multiple reports of reversion of invasive to benign epithelial tumour phenotype following ectopic expression of E-cadherin in carcinoma cells, establishing the tumour suppressive role of E-cadherin (Frixen *et al.*, 1991; Birchmeier and Behrens 1994; Bracke *et al.*, 1996). Seminal work by Perl and colleagues in 1998 demonstrated a causal role of E-cadherin loss for the transition from well-differentiated to invasive carcinoma in a transgenic mice model, Rip1Tag2, for development of pancreatic cancer (Perl *et al.*, 1998). Notably, tumorigenesis is arrested at the adenoma stage in cells preserving E-cadherin expression, whereas

expression of a dominant negative mutant of E-cadherin instigated early invasion and metastasis, indicating loss of E-cadherin as a rate-limiting step in carcinogenesis.

Oda and colleagues undertook one of the first efforts to decipher molecular events leading to loss of E-cadherin expression (Oda *et al.*, 1994). They analysed 10 human cancer cell lines which exhibited loose cell-cell adhesion and this early work already hinted at the multiple mechanisms by which E-cadherin can be dysregulated. Out of the 10 human cell lines, 4 of them expressed high level of wildtype E-cadherin mRNA and protein, suggesting an aberration in E-cadherin regulation or in other components of cell adhesion system. (Indeed, mutations in α-catenin (Shimoyama *et al.*, 1992; Oda *et al.*, 1993) and β-catenin (Oyama *et al.*, 1994) have been implicated in disruption of cell adhesion). Another 4 did not express E-cadherin, indicating the occurence of gene silencing event. In the other 2, mutations were found at exon-intron junctions, which resulted in aberrant splicing and presence of 4 transcripts with different inserts. By genetic profiling of familial gastric cancer, Guilford *et al.* identified germline mutations in E-cadherin, which were responsible for diffuse-type gastric cancer (Guilford *et al.*, 1998). This firmly established the key role of E-cadherin mutations in cancer development (Birchmeier 2005).

In addition to genetic aberration, loss of E-cadherin expression also happens as a consequence of epigenetic alterations (Baranwal and Alahari, 2009). Yoshiura *et al.* revealed that the promoter region of the E-cadherin gene is hypermethylated around CpG islands in some E-cadherin-negative cancer cell lines, thereby resulting in the silencing of its expression (Yoshiura *et al.*, 1995). Such a mechanism of silencing has also been observed with the tumour suppressors, Rb (Greger *et al.*, 1989; Sakai *et al.*, 1991) and von hippel lindau (VHL) (Herman *et al.*, 1994). Transcriptional repression represents another mode of epigenetic regulation that can suppress E-cadherin expression (see section 1.1.2.1.1). Multiple transcriptional repressors have already been implicated in cancer development (van Roy and Berx, 2008). A strong correlation was obtained between overexpression of Snail transcription factors and reduced E-cadherin expression in invasive ductal breast carcinoma (Cheng *et al.*, 2001). In line with this, inhibition of Snail function in E-

cadherin-negative epithelial tumour cells restores E-cadherin expression (Battle *et al.*, 2000). Different gastric cancers subtypes have also been shown to be associated with different E-cadherin transcription repressors (Rosivatz *et al.*, 2002). For diffuse gastric cancers, arising due to inactivating mutation of E-cadherin, a strong correlation was observed between Snail upregulation and E-cadherin loss. Interestingly, Snail is not upregulated in intestinal type gastric cancer. Instead, these gastric cancer samples are associated with high expression level of SIP1/ZEB2. The differential expression level of the different transcription repressors in different gastric cancer subtypes therefore suggests that they are not functionally redundant and hence play different roles in tumour development. Twist, a basic helix-loop-helix (bHLH) transcription factor, which is capable of repressing E-cadherin expression, was found highly expressed in metastatic breast cancer samples (Yang *et al.*, 2004).

It is, perhaps, no coincidence that these transcription factors are also key players in inducing the transcriptional program required for epithelial-to-mesenchymal transition (EMT) (Huber *et al.*, 2005; Yang and Weinberg, 2008). Morphologically, as its name implies, EMT refers to a cellular process by which epithelial cells undergo transformation to shed epithelial features and assume mesenchymal characteristics, such as enhanced migratory capacity, invasiveness and heightened resistance to apoptosis (Kalluri and Neilson, 2003). The concept of EMT was firmly established by a series of elegant experiments carried out by Greenburg and Hay, which demonstrated that epithelial cells in culture may acquire mesenchymal characteristics (Greenburg and Hay, 1982; Greenburg and Hay, 1986; Greenburg and Hay, 1988). It is a cellular programme that is induced at several morphogenetic stages of embryonic development and during wound healing (Weinberg, 2007). It is widely accepted now that this physiological programme is hijacked by cancer cells to break away from cell-cell adhesion and acquire motility, i.e. to become metastatic (Thiery and Sleeman, 2006; Hugo *et al.*, 2007). In agreement with the E-cadherin-repression role of these transcriptional factors, loss of E-cadherin is considered a hallmark of EMT (Weinberg, 2007; Yang and Weinberg, 2008). Stoker and Perryman provided the first evidence that EMT can be elicited by growth factor signalling (Stoker and Perryman, 1985). They discovered that

incubation of polarised MDCK epithelial cells with conditioned fibroblast medium induced scattering of the cells, whereby the cells were detached from each other and acquired fibroblastic features, and subsequently isolated the “scatter factor”, which is also known as hepatocyte growth factor (HGF). This was followed by discovery of transforming growth factor-β (TGF-β), a serine/threonine receptor kinase, as a potent inducer of E-cadherin repression and EMT in mammary epithelial cells (Miettinen *et al.*, 1994). With the establishment of EMT as a key process in tumour progression, the immense research interest drawn towards this cellular program has elucidated other signalling pathways that are capable of eliciting EMT, such as canonical Wnt signalling, Hedgehog signalling, Notch signalling, epidermal growth factor receptor (Siegfried *et al.*, 1992), signalling, platelet-derived growth factor receptor (PDGFR) signalling and nuclear factor kappa-light-chain-enhancer of activated B cell (NF-κB) signalling (Huber *et al.*, 2005; Thiery, 2006). These pathways can be activated by the presence of their specific ligands or regulated by crosstalk among each other. Notably, these pathways have been shown to induce EMT via the E-cadherin repressors, Snail, Slug, SIP1/ZEB2 and Twist (Figure 1.6) (Huber *et al.*, 2005; Yang and Weinberg, 2008). Therefore, the aberrant signalling of these pathways can all lead to loss of E-cadherin expression.

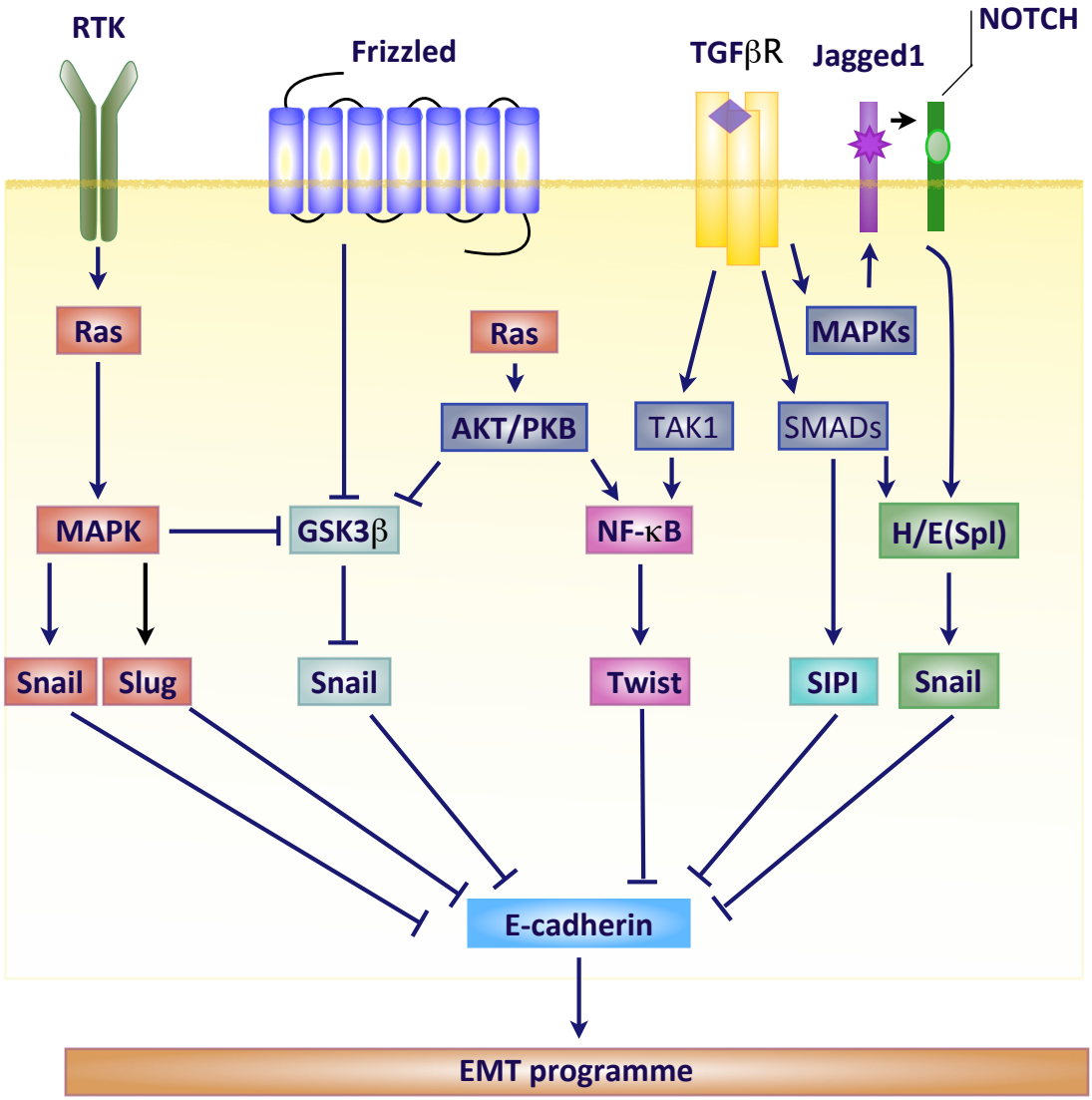

Figure 1.6. Silencing of E-cadherin expression downstream of various signalling pathway involved in epithelial-to-mesenchymal transition (EMT). Activation of receptor tyrosine kinase (RTK), Wnt signalling, TGFβR and Notch signalling pathways induces the transcriptional repressors, including Snail, Slug, Twist and SIP1, which repress E-cadherin expression. The loss of E-cadherin is considered a hallmark and a pre-requisite for induction of a full EMT program.

Beside the genetic and epigenetic mechanisms by which E-cadherin expression is silenced, dysregulation at the post-translational level can also contribute to disruption of normal E-cadherin function in pathological conditions such as cancer (van Roy and Berx, 2008). The activation of receptor tyrosine kinases, such as HGF and the subsequent activation of its downstream effector Src, results in enhanced tyrosine phosphorylation on the cytoplasmic tail of E-cadherin (Fujita *et*

al., 2002). This is followed by recruitment of Hakai, a ubiquitin ligase, which specifically interacts with and ubiquitylates the cytoplasmic tail of phosphorylated E-cadherin, thereby promoting its internalisation and subsequent clearance via the lysosomal pathway. Besides, the stability of AJ complex is also determined by its cytoplasmic tail binding partner – p120-catenin and β-catenin. Reynolds' laboratory has demonstrated that p120-catenin plays a crucial role in regulating E-cadherin stability, where the siRNA depletion of p120-catenin is accompanied by significant loss of E-cadherin protein, but not mRNA (Davis *et al.*, 2003). Downregulation of p120-catenin has been reported in multiple cancer types, including colon, prostate, breast and lung carcinoma (Thoreson and Reynolds, 2002; Reynolds and Carnahan, 2004). Therefore, it is possible that, at least for a subset of these cancers, the E-cadherin deficiency occurs as a result of the decreased level of p120-catenin. In addition, cleavage of E-cadherin ectodomain by matrix metalloproteases (MMP) or the serine proteases kallikrein, represents another way through which E-cadherin function is inhibited (De Wever *et al.*, 2007; van Roy and Berx, 2008). For example, overexpression of kallikrein 6 (klk6) in pancreatic adenocarcinomas enhances release of soluble E-cadherin fragment (Klucky *et al.*, 2007), which may functions as pseudoligands that interfere with normal E-cadherin interaction, thereby promoting invasion (De Wever *et al.*, 2007; Klucky *et al.*, 2007).

1.1.2.2 β-catenin

β-catenin was first identified in *Drosophila* as a product of the segment polarity gene Armadillo, which when mutated, resulted in deletion of a defined fraction of the pattern in each segment (Nusslein-Volhard and Wieschaus, 1980). Subsequently, together with its close relative plakoglobin, β-catenin was shown to be component of adherens junctions and form a direct interaction with E-cadherin (Ozawa *et al.*, 1989; McCrea *et al.*, 1991). β-catenin is a 781 aa long protein, with approximately 130 aa on its N-terminus, followed by a central region of about 550aa that contains 12 armadillo repeats and another 100 aa on its C-terminus (Figure 1.7A) (Huber *et al.*, 1997). The armadillo repeat represents an imperfect repeat of a

42 aa motif, which was originally defined in the *Drosophila* segment polarity gene Armadillo (Riggleman *et al.*, 1989) and is capable of forming a triple α-helix (Huber *et al.*, 1997). The 12 repeating triple α-helices altogether form a right-handed α-solenoid protein structure, where the polypeptide chain forms a continuous superhelix (Figure 1.7B) (Kobe and Kajava, 2000). The armadillo superhelix forms a positively charged shallow groove and allows for binding with multiple interacting partners, such as E-cadherin (Hulsken *et al.*, 1994), the scaffolding protein axin (Xing *et al.*, 2003), T-cell factor/Lymphoid enhancer factor (TCF/LEF) family transcription factors (Behrens *et al.*, 1996) and adenomatous polyposis coli (APC) tumour suppressor protein (Hulsken *et al.*, 1994; Rubinfeld *et al.*, 1995). Notably, the binding of β-catenin with the different partners is mutually exclusive and hence the binding partner dictates the function of β-catenin (Hulsken *et al.*, 1994; Daugherty and Gottardi, 2007). At the N-terminus of β-catenin, there is a cluster of 4 serine (S)/threonine (T) residues (S33, S37, T41 and S45), which are conserved from *Drosophila* to human (Peifer *et al.*, 1994). The sequential phosphorylation of these S/T residues, first a priming phosphorylation of S45 by CK1, followed by phosphorylation of S33, S37 and T41 by GSK3β, alters the stability of β-catenin (Liu *et al.*, 2002). The phosphorylated S33 and S37 residues allow recognition of β-catenin by the β-transduction repeat containing protein (βTrCP) E3-ligase, which ubiquitylates β-catenin and earmarks it for proteasomal degradation (Hart *et al.*, 1999). The C-terminus of β-catenin represents a transactivation domain, which allows transcriptional activation when β-catenin is bound to TCF/LEF1 transcription factors (Willert and Nusse, 1998).

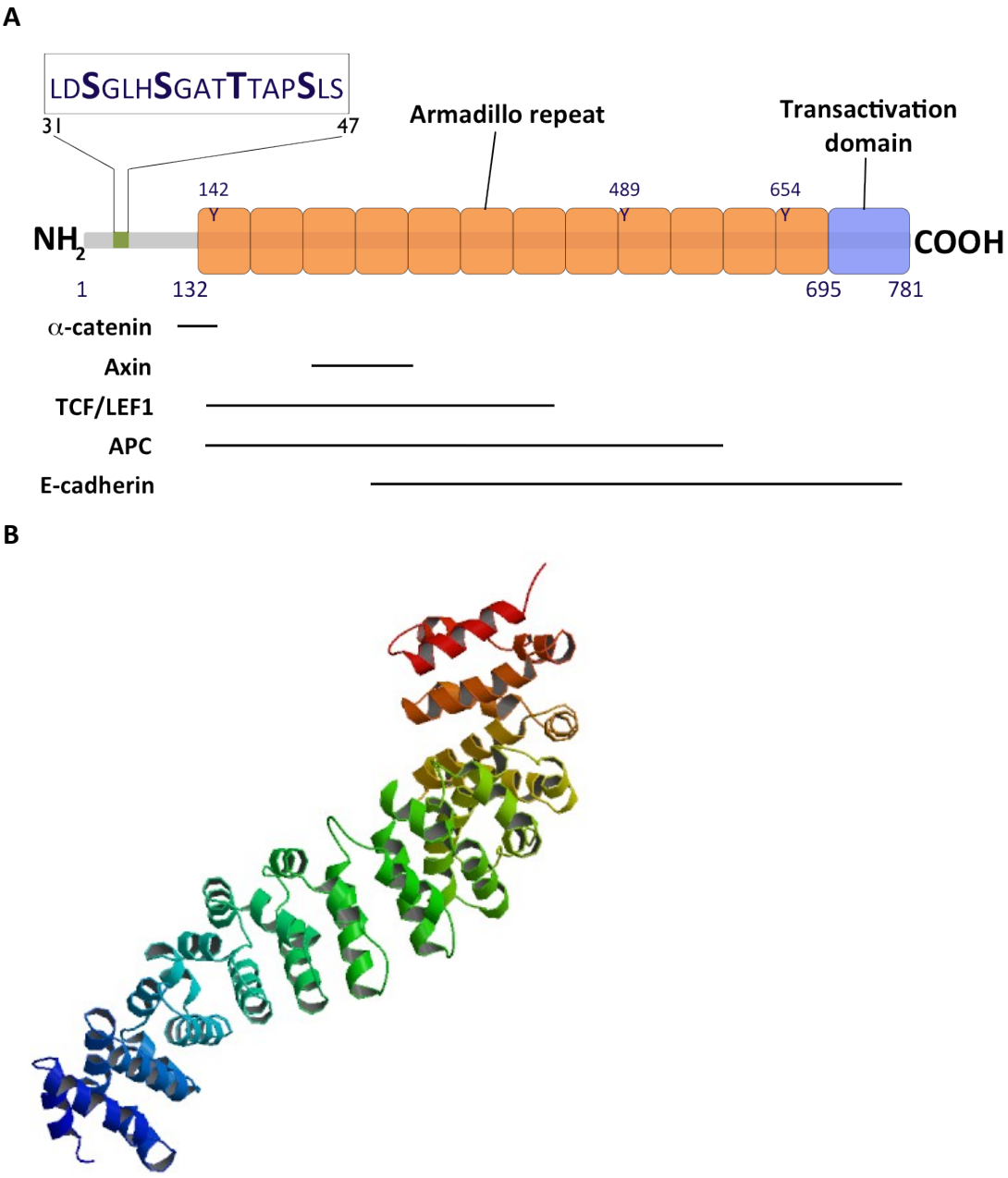

Figure 1.7. β-catenin structure. (a) Domain architecture of β-catenin. β-catenin is characterised by the 12-armadillo repeats at the centre of the protein. Tyrosine residues (Y142, Y489 and Y654) in the armadillo repeats which disrupts E-cadherin-β-catenin interaction upon phosphorylation are shown. The highly conserved consensus motif for GSK3β phosphorylation near the N-terminal end is shown, with the Serine/Threonine phosphorylation sites highlighted in bold. The C-terminus of β-catenin is involved in transcriptional activation upon binding with TCF/LEF1. The binding sites for different β-catenin interacting partners, α-catenin, axin, TCF/LEF1, APC and E-cadherin are marked by the dark lines. **(b) Crystal structure of armadillo repeat of β-catenin.** Each armadillo repeat folds into a triple-helix structure and the 12 armadillo repeats of together form a superhelix, with a shallow grove to allow binding of the β-catenin interacting partners. The structure was retrieved from Research Collaboratory for Structural Bioinformatics Protein Data Bank (PBD ID = 2BCT).

1.1.2.2.1 Cell adhesion function of β-catenin

As described at the beginning of Section 1.1.2, β-catenin is a core component of the AJ. It associates with E-cadherin at the ER and this interaction is a prerequisite for subsequent transport of the E-cadherin-β-catenin complex to the plasma membrane (Chen *et al.*, 1999). At the AJ, it provides a link between E-cadherin and α-catenin, and its stable association with E-cadherin at the plasma membrane is also important for the maintenance of cell adhesion (Gloushankova, 2008). Amino acid residues 625-723 of mature E-cadherin interact across the entire superhelix consisting of 12 Armadillo repeats (Huber and Weis, 2001). As suggested by Huber and Weis, this extended interface allows "multiple, quasi-independent interactions" that "can be regulated independently, enabling combinatorial regulation of the interaction and the integration of multiple input signals".

The interaction between E-cadherin and β-catenin is very much dependent upon the phosphorylation status of the two proteins (Daugherty and Gottardi, 2007). *In vitro* phosphorylation of the cytoplasmic tail of E-cadherin at S838, S840 and S846 generates more molecular contacts on the binding surfaces, thereby enhancing the affinity of β-catenin for E-cadherin by 300-fold (Huber and Weis, 2001; Choi *et al.*, 2006). However, the *in vivo* kinases responsible for such phosphorylation remain to be elucidated (Daugherty and Gottardi, 2007). On the other hand, phosphorylation of S844 by Casein Kinase 1 (CK1) weakens β-catenin binding to E-cadherin and promotes E-cadherin internalisation (Dupre-Crochet *et al.*, 2007).

There are 3 conserved tyrosine (Y) residues in β-catenin at the E-cadherin-β-catenin interface, namely Y142, Y489 and Y654 (Lilien and Balsamo, 2005). Phospho-regulation of these residues represents a mechanism by which integrity of the AJ complex is regulated. Y142 is located within the α-catenin binding region of β-catenin (aa 118 – 149) and its phosphorylation disrupts interaction between β-catenin and α-catenin, followed by loss of cell-cell adhesion (Piedra *et al.*, 2003). The tyrosine kinases targeting Y142 include c-Met (Brembeck *et al.*, 2004), Fyn and

activated Fer kinases (Piedra *et al.*, 2003). Phosphorylation of Y654, in the 12th Armadillo repeat, by EGFR or Src resulted in a 6-fold decrease in affinity with which β-catenin binds E-cadherin (Huber and Weis, 2001). This is because the phosphorylated Y654 clashes with a key aspartate residue on E-cadherin. Phosphorylation of Y489 by the Abelson (Abl) kinase has been shown to disrupt N-cadherin/β-catenin interaction (Rhee *et al.*, 2002). Therefore, phosphorylation of Y489 may disrupt E-cadherin/β-catenin association as well.

1.1.2.2.2 β-catenin – at the heart of canonical Wnt signalling

β-catenin was first recognised for its cell adhesive function (Ozawa *et al.*, 1989). The first hint of its signalling function came only when it was identified as a mammalian homologue of armadillo, a segment polarity gene involved in the wingless signalling pathway, which is crucial for development of *Drosophila* (McCrea *et al.*, 1991). Mutations of armadillo and other genes including wingless, dishevelled, shaggy/zeste-white 3 (homologue of GSK3β) all resulted in cuticle abnormalities in the embryo of *Drosophila* (Nusslein-Volhard and Wieschaus, 1980). Multiple epistatic analyses carried out to assess cuticle structure in double mutants of these genes revealed that these genes are core components of a new signal transduction pathway – Wnt signalling (Siegfried *et al.*, 1992; Noordermeer *et al.*, 1994; Peifer *et al.*, 1994). The Wnt signalling pathway plays fundamental roles in regulating cell proliferation, cell polarity and determining cell fate during cellular processes such as metazoan development, tissue homeostasis and cancer (Clevers, 2006; Macdonald *et al.*, 2009; Mosimann *et al.*, 2009).

In human, there are 12 subfamilies of secreted Wnt proteins and a total of 20 members (Clevers, 2006). There are three different pathways known to be triggered upon Wnt receptor activation, namely the canonical Wnt signalling pathway, non-canonical planar cell polarity (PCP) pathway and the Wnt/Ca^{2+} pathway, among which the former is most well-characterised. β-catenin is the key effector of the canonical Wnt signalling pathway and regulation of its expression level is central to

the signalling events triggered upon activation of the pathway (Morin, 1999). In the absence of Wnt ligand, β-catenin is continually synthesised and is bound to a "destruction complex" consisting of axin, APC, GSK3β and CK1 (reviewed in Clevers, 2006; Mosimann, 2009). In this complex, Axin is the scaffolding protein with specific binding sites for β-catenin, CK1, GSK3β and APC (Daugherty and Gottardi, 2007). Upon binding of β-catenin to the "destruction complex", CK1 first phosphorylates β-catenin at its S45 and this priming phosphorylation then enables subsequent phosphorylation of S33, S37 and T41 by GSK3β (Liu *et al.*, 2002). The N-terminally phosphorylated β-catenin is then recognised by the βTrCP E3 ligase, which ubiquitylates it and targets it for proteasomal degradation (Hart *et al.*, 1999).

When Wnt ligand is present and binds to the seven transmembrane domain receptor Frizzled and the co-receptor, low density lipoprotein receptor-related protein 5 or 6 (LRP5/6), the LRP5/6 becomes phosphorylated by CK1 and GSK3β (Figure 1.8) (Davidson *et al.*, 2005; Zeng *et al.*, 2005) and the cytoplasmic dishevelled (Dvl) is recruited to Frizzled, where it polymerises with other Dvl (Bilic *et al.*, 2007; Schwarz-Romond *et al.*, 2007). The phosphorylation of LRP5/6 and polymerisation of Dvl at the receptor complex allows efficient recruitment of Axin, thereby preventing assembly of the destruction complex. As a result, β-catenin is not phosphorylated and ubiquitylated, therefore not degraded. The accumulating hypophosphorylated form of β-catenin then translocates into the nucleus (Henderson and Fagotto, 2002) and interacts with the TCF/LEF1 family proteins to form a bipartite transcription complex, which interacts with other co-activators such as BCL9, Pygopus and CREB-binding protein (CBP) (reviewed in Klaus and Birchmeier, 2008) and initiates transcription of β-catenin target genes (Behrens, 1996; Mosimann *et al.*, 2009).

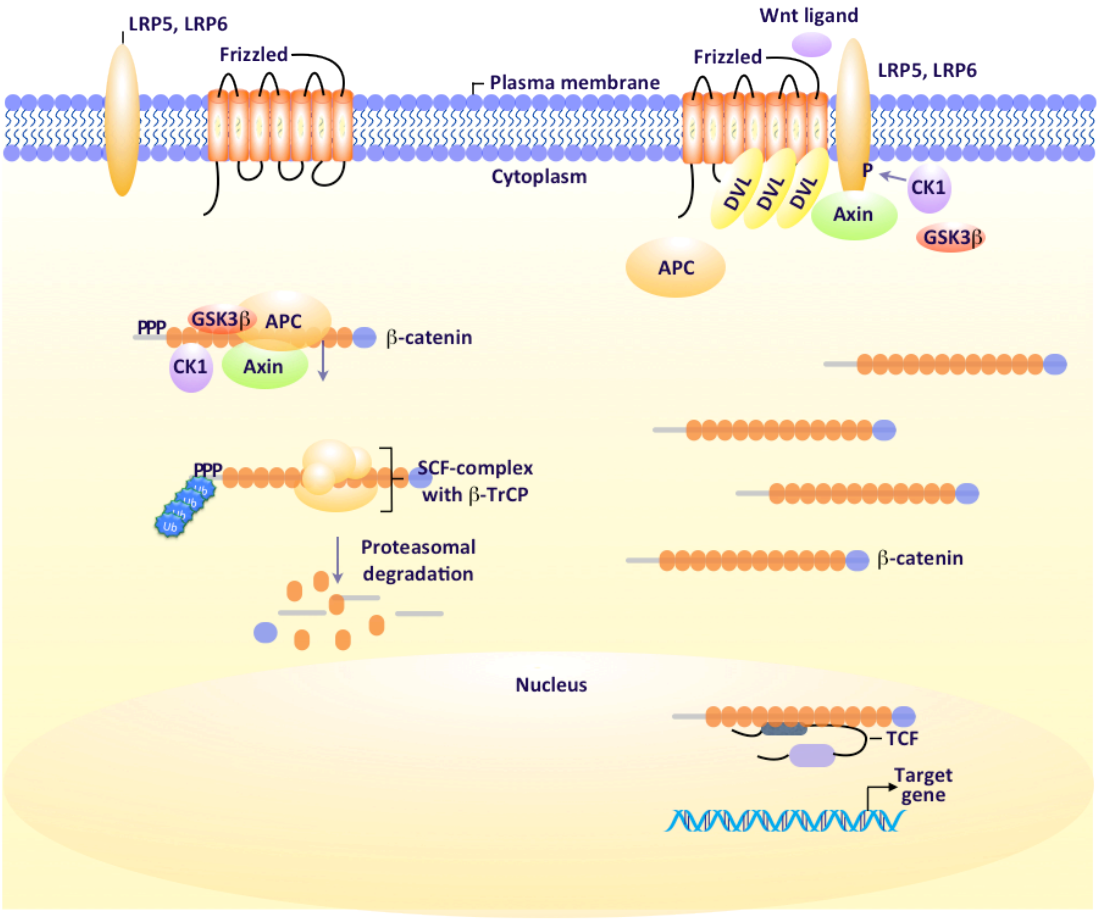

Figure 1.8. Wnt signalling pathway. In the absence of Wnt ligands, β-catenin is bound by the destruction complex consisting of GSK3β, Casein kinase 1 (CK1), APC and Axin. The GSK3β and CK1 phosphorylates β-catenin on its N-terminus, and the phosphorylated β-catenin is then targeted by the SCF-βTrCP ubiquitin ligase complex for polyubiquitylation and is subjected to proteasomal degradation. In the presence of Wnt ligand, the desctruction complex is disassembled due to sequestration of Axin to the plasma membrane, which is partly mediated by the multimerised-Dishevelled associated with the Frizzled receptor. β-catenin is therefore not phosphorylated and not targeted for polyubiquitylation and proteasomal degradation, and becomes stabilised in the cytoplasm. The accumulating hypophosphorylated β-catenin then translocates to the nucleus and interacts with TCF/LEF1, forming a transcriptional complex which induces expression of Wnt target genes.

1.1.2.2.3 β-catenin/Wnt Signaling and cancer

Aberrant regulation of cell signalling pathway, normally involved in control of cell proliferation, motility and survival, has become a common mechanism by which cancers develop (Martin 2003). Wnt signalling is no exception to this general theme.

The first link of Wnt signalling and cancer was demonstrated when the tumour suppressor gene *APC*, first found mutated in familial adenomatous polyposis (FAP) (Groden *et al.*, 1991; Nishisho *et al.*, 1991), was found interacting with β-catenin (Rubinfeld *et al.*, 1993). As mentioned in the previous section, APC is a core component of the β-catenin destruction complex. The mutational inactivation of this gene therefore leads to inappropriate stabilisation of β-catenin (Rubinfeld *et al.*, 1996), its translocation to the nucleus and subsequent activation of the genetic program elicited by the β-catenin/TCF transcription complex (Reya and Clevers, 2005; Brembeck *et al.*, 2006). In fact, nuclear accumulation of β-catenin is regarded as the hallmark of canonical Wnt pathway activation (Giles *et al.*, 2003) and persistence of β-catenin/TCF complex forms the basis of tumour transformation of epithelial cells (Clevers 2004; Reya and Clevers, 2005).

Multiple components of the Wnt signalling pathway are found mutated in cancers (Polakis 2000; Brembeck *et al.*, 2006). Regardless of the nature of the mutations or identity of the mutated components, the end effect is β-catenin stabilisation. In addition to mutations in hereditary cancer like FAP, APC is also found mutated in 80% of sporadic colorectal cancers (Kinzler and Vogelstein, 1996). Most of these mutations result in truncation of APC, which diminishes its ability to bind β-catenin or Axin and hence β-catenin is stabilised (Polakis, 2000). However, mutation of APC alone is not sufficient to drive tumour progression, but additional mutations in other pathway components, such as oncogenic activation of Ras and loss-of-function mutation of p53, are also required (Vogelstein and Kinzler, 1993; Vogelstein and Kinzler, 2004). Mutations in the N-terminal regulatory region of β-catenin are also responsible for activation of Wnt signalling pathway in a wide range of cancers, including colon cancer, melanoma, prostate cancer, hepatocellular carcinoma, ovarian cancer, endometrial carcinomas, medulloblastomas and pilomatricomas (reviewed in Morin, 1999). Figure 1.9 summarises the mutations that have been identified in regulatory region of β-catenin in a range of cancers. As discussed in section 1.1.2.2, this regulatory region contains 4 conserved serine/threonine residues, the phosphorylation of which, by CK1 and GSK3β, is important for

subsequent polyubiquitylation by β-TrCP E3 ligase and degradation (Liu *et al.*, 2002). In line with this experimental evidence, these conserved serine/threonine residues in the regulatory region of β-catenin are most frequently mutated (Morin, 1999). For example, in 50% of colon tumors with wildtype APC, mutations were found in the regulatory region of β-catenin (Sparks *et al.*, 1998). In two APC-positive colon cancer cell lines harbouring mutations in the β-catenin regulatory region, namely HCT116 (deletion of S45) and SW48 (S33T mutation), the mutant β-catenin is not regulated by APC and the cell lines exhibit increased β-catenin/TCF transcriptional activity as measured by a reporter assay (Morin *et al.*, 1997). Mutations were also found at D32 and G34, both of which are next to the conserved phosphorylation residue S33. These mutations possibly results in conformational changes around the S33 residue, which prevents phosphorylation or recognition by the β-TrCP E3-ligase, as this residue is part of a ubiquitin-dependent recognition sequence (Aberle *et al.*, 1997; Orford; *et al.*, 1997). Biallelic inactivation of Axin leading to truncated mutants of the proteins was detected in hepatocellular carcinoma and cell lines (Satoh *et al.*, 2000). These mutants were predicted to lack β-catenin binding activities and therefore not able to target β-catenin for ubiquitin-dependent proteasomal degradation.

In addition to mutations, changes in the expression level of the different components may contribute to aberrant Wnt signalling as well (Barker and Clevers, 2006). For example, overexpression of Wnt ligands and Frizzled receptors have been reported in colon cancer (Holcombe *et al.*, 2002), breast cancer (Milovanovic *et al.*, 2004) and head and neck cancer (Rhee *et al.*, 2002). Increased expression of dishevelled family members has also been associated with mesothelioma (Uematsu *et al.*, 2003), non-small-cell lung cancer (Uematsu *et al.*, 2003) and cervical cancer (Vogelstein and Kinzler, 1993). On the other hand, reduced expression of the natural Wnt inhibitors of the SFRP (secreted Frizzled-related protein) and WIF (Wnt inhibitor factor) family have also been reported in colon cancer (Caldwell *et al.*, 2004; Suzuki *et al.*, 2004), breast cancer (Ugolini *et al.*, 2001; Ai *et al.*, 2006), and mesothelioma (Lee *et al.*, 2004; Batra *et al.*, 2006).

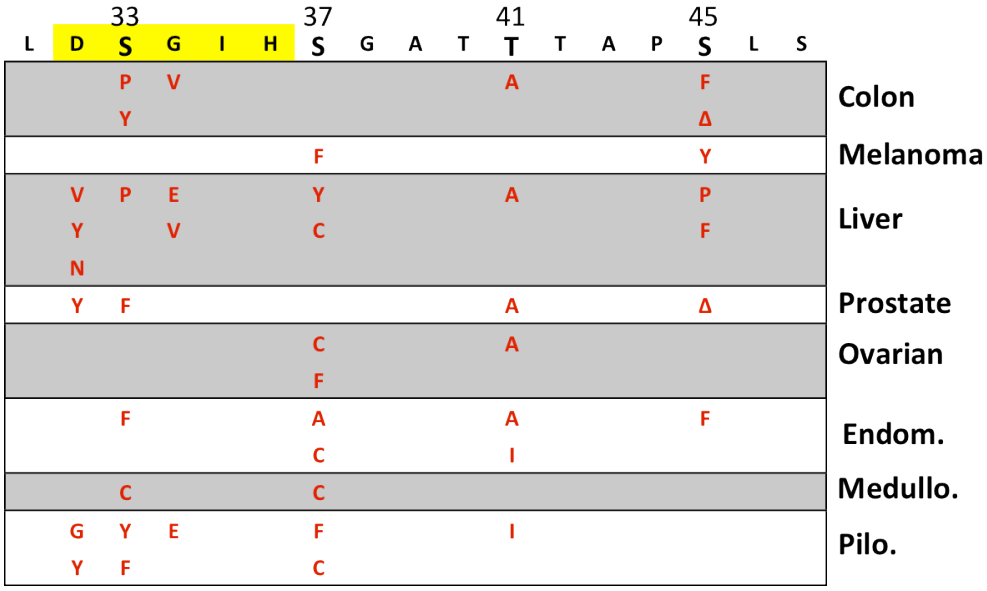

Figure 1.9. Cancer associated mutations found in the N-terminal regulatory region of β-catenin. The conserved phosphorylation Serine/Threonine residues are highlighted in bold, and the red letters below are the residues found after mutation in cancer. Region highlighted in yellow is the ubiquitylation motif. (Δ = deletion mutation; Endom. = endometrial carcinoma; Medullo. = medulloblastomas; Pilo. = pilomatricomas) (reviewed in Morin, 1999).

The research efforts around Wnt signalling pathway has provided substantial understanding of the molecular mechanisms regulating the pathway and therefore allowed design of therapeutic strategies to treat cancer with activated Wnt signalling (Barker and Clevers, 2006; Klaus and Birchmeier, 2008). For example, in the case of overexpressed Wnt and Frizzled, where there is no mutation among components of the signalling pathway, antibodies against the two proteins represent a potential cancer therapeutic (You *et al.*, 2006). Indeed, by treating a head and neck cancer cell line with overexpressed Wnt1 with a Wnt1 antibody, Rhee and colleagues demonstrated that Wnt signalling was suppressed in the cell, and the cells stopped proliferating and underwent apoptosis (Rhee *et al.*, 2002). Small-molecule inhibitors, which can specifically disrupt formation of β-catenin-TCF complex, also hold great promise for treatment of Wnt signalling activated cancer (Barker and Clevers, 2006). However, efforts remain to be made in the search for safe and effective drugs to block Wnt signalling activity in cancers.

1.2 Post-translational modification by ubiquitin

1.2.1 Emergence of the concept of protein degradation

It is now a fundamental understanding of cell biology that all proteins have a definite life-span and are all eventually degraded. However, up until the early decades of the 20th century, the general view was that intracellular proteins, whether structural or functional, represent stable constituents of the body whereas dietary proteins serve the sole purpose of energy provision (Ciechanover, 2005). With a new tool – heavy isotope labeling – becoming available, Schoenheimer and Rittenberg decided to challenge this paradigm (Schoenheimer *et al.*, 1939). By feeding rat with ^{15}N-labelled tyrosine, they discovered that only 50% of the heavily labeled amino acid was recovered in the urine while the remainder was incorporated into tissue proteins, and this was accompanied by excretion of an equivalent amount of protein nitrogen. Such observation demonstrated unequivocally that intracellular proteins are dynamic entities, which are continually synthesized and degraded. As unambiguous as the experimental result was, the concept of protein turnover did not receive wide acceptance and was still challenged until mid 1950s.

Another significant milestone in establishing the concept of protein turnover was finally laid when Christian de Duve and his colleagues identified lysosomes, elucidating the first cellular machinery that mediates intracellular proteolysis (De Duve *et al.*, 1953). Lysosome is a membrane-bound vacuolar structure with hydrolytic enzymes bathed in an acidic internal environment and is the site where engulfed macromolecules, endocytosed membrane proteins, worn-out organelles and certain intracellular proteins are degraded (Pillay *et al.*, 2002). However, accumulating evidence in the next two decades hinted at existence of a non-lysosomal proteolytic pathway (Ciechanover, 2005). Firstly, it was revealed that the degradation of proteins in bulk mediated by lysosome is non-selective and it was conceptually difficult to reconcile this observation with the emerging concept that the turnover rate of proteins varies from one to the other (Schimke and Doyle, 1970; Goldberg and St John, 1976). Secondly, inhibition of lysosomal proteases using

specific inhibitors had differential effects on different populations of proteins, suggesting existence of an alternative degradative mechanism. The key experiment leading to this proposal was performed by Poole and colleagues (Poole *et al.*, 1977). They demonstrated that degradation of extracellular proteins, but not of intracellular proteins, was sensitive to inhibition of lysosomal function by weak bases. Thirdly, while it was known that the proteolytic actions carried out by lysosomal proteases occur exergonically, there was mounting evidence indicating the energy requirement of degradation of cellular proteins.

These findings have prompted the quest for a proteolytic mechanism alternative to lysosome that exhibits high degree of specificity towards its substrates.

1.2.2 Discovery of ubiquitin proteasome system

A major challenge in identifying a non-lysosomal proteolytic system lay in the development of an *in vitro* biochemical assay that can faithfully recapitulate the non-lysosomal cellular proteolytic events (Ciechanover, 2005). This was made possible by the important discovery of Rabinovitz and Fisher, who found that mature rabbit reticulocytes which lack lysosomes are capable of degrading abnormal haemoglobin, providing a lysosome-free cell-based system for investigation of an alternative proteolytic system (Rabinovits and Fisher, 1964). Following this, Etlinger and Goldberg was the first to prepare a crude extract from reticulocytes, which preserves proteolytic activity against abnormal haemoglobin, in an adenosine triphosphate (ATP)-dependent manner and at a neutral pH (indicating its non-lysosomal origin) (Etlinger and Goldberg, 1977). This crude extract from reticulocytes proved to be instrumental in the efforts to identify the alternative proteolytic system.

Ciechanover and Hershko demonstrated that the crude extract could be fractionated on an anion-exchange resin into 2 fractions, fractions I and II, and that both of them were required for reconstitution of proteolytic activity (Ciechanover *et al.*, 1978). They identified a ~8.5kDa heat-stable protein in fraction I, which was

thought to be an activator for a protease in fraction II. This active component was later demonstrated to be covalently conjugated to target substrate in the presence of fraction II in an ATP-dependent manner, and was then named ATP-dependent proteolysis factor 1 (APF1) (Ciechanover *et al.*, 1980; Hershko *et al.*, 1980). A model of protein degradation was then proposed: covalent conjugation of several moieties of APF1 to a protein substrate targets it for degradation by a yet-to-be-identified protease which does not recognize the unmodified protein. Independent to these findings, a protein named "ubiquitin" (due to its ubiquitous expression) was found to be covalently conjugated to histones H2A and H2B (Goldknopf and Busch, 1975; Goldknopf and Busch, 1977; Hunt and Dayhoff, 1977), in a similar fashion to how APF1 is conjugated to target substrates. This raised the suspicion that APF1 and ubiquitin could be the same protein and sequence analysis revealed that was indeed the case (Wilkinson *et al.*, 1980). Using immobilized ubiquitin as a 'covalent' affinity bait, Hershko, Ciechanover and colleagues purified the three enzymes which act in a sequential manner to catalyse ubiquitin conjugation. The enzymes are namely E1, the ubiquitin activating enzyme, E2, the ubiquitin-conjugating enzyme and E3, the ubiquitin ligase (Ciechanover *et al.*, 1982; Hershko *et al.*, 1983).

The last piece of the puzzle however remained to fall in place: the identity of the downstream protease in fraction II that recognises ubiquitin-modified substrates remained elusive. It was known that there was a second ATP-requiring step in the proteolytic system following ubiquitin conjugation based on work by Tanaka and colleagues (Tanaka *et al.*, 1983), and Hershko and colleagues subsequently demonstrated that degradation of ubiquitin-modified proteins requires energy (Hershko *et al.*, 1984). An important breakthrough was made by Hough and colleagues, when they purified a high-molecular-mass proteolytic complex that degraded ubiquitylated lysozyme in an ATP-dependent manner (Hough *et al.*, 1986). The proteolytic complex fitted all the criteria for a ubiquitin-regulated degradative machinery and was later named proteasome.

The eukaryotic 26S proteasome is a barrel-like protein complex, comprising of two subcomplexes: a 20S proteolytic core particle and one or two 19S regulatory

particles (Figure 1.10) (Groll and Clausen, 2003; Pickart and Cohen, 2004). The core particles consists of 28 subunits, which are arranged into four seven-membered rings that stack on top of each other. The two middle rings are made up of β-subunits only and together they form the catalytic chamber, whereas the two outer rings consists of α-subunits only and serve as a narrow gate to control entry and exit of polypeptides and degradative products respectively. The 19S regulatory particle consists of a base, which is made up of six ATPase and two non-ATPase proteins; and a lid, which is composed of eight non-ATPase proteins. The base and lid, of human proteasomes, are linked by a subunit called S5a. The lid of the regulatory particle plays a dual role in recognition of polyubiquitylated substrate and unfolding of substrate. The S5a, Rad23 and Dsk2 subunits of the regulatory particle have all been shown to be able to bind and sense the length of polyubiquitin chains attached to substrate (Elsasser *et al.*, 2004; Zhang *et al.*, 2009), whereas the unfolding of protein into polypeptide chains is mediated by the activities of the ATPases (Pickart and Cohen, 2004). Once a polypeptide chain enters the chamber of the core particle, it is degraded non-specifically.

In summary, the series of findings discussed herein had led to the identification of the core components of a non-lysosomal proteolytic system: a tag for specific degradation of protein, the ubiquitin and an executer of proteolysis, the proteasome, hence the ubiquitin proteasome system (UPS). Perhaps what was unexpected and only became apparent after more than a decade following the discovery of UPS, is that regulated lysosomal degradation of endocytosed membrane protein requires modification of the substrate protein by ubiquitin as well (Hurley and Stenmark, 2011).

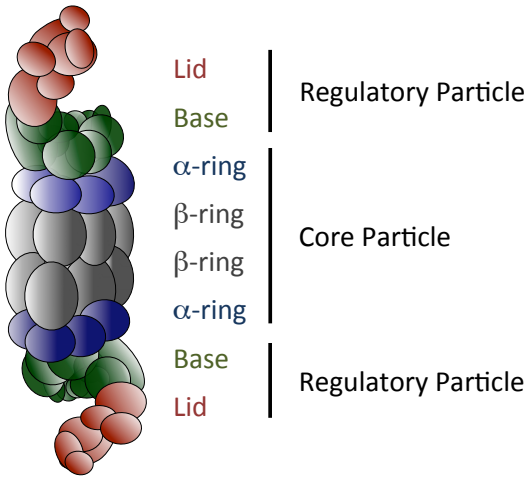

Figure 1.10. The structure of the eukaryotic 26S proteasome. The eukaryotic 26S proteasome consists of a 20S proteolytic complex, which is the core particle and one or two 19S complexes, which is the regulatory particle. The core particle consists of four seven-membered rings stacked on top of each other: two outer a-rings and two inner b-rings. The 19S regulatory particle consists of a base, which is made up of six ATPase and two non-ATPase proteins; and a lid, which is composed of eight non-ATPase proteins. The base and lid, of human proteasomes, are linked by a subunit called S5a.

1.2.3 Ubiquitin and Reversible Ubiquitylation

Ubiquitin is a globular protein of 76 amino acids that is highly conserved among all eukaryotes (Hershko and Ciechanover, 1998). It was first isolated from samples of bovine thymus and was named ubiquitin owing to its ubiquitous expression across species (Goldstein *et al.*, 1975). Ubiquitin is encoded by four different genes, namely UBB, UBC, UBA52 and RB27A in human. Translation of UBB and UBC mRNA gives rise to polyproteins consisting of tandem repeats of nine ubiquitin (Wiborg *et al.*, 1985), whereas UBA52 and RB27A give rise to ubiquitin fused N-terminally to the ribosomal proteins, L40 and S27a, respectively (Baker and Board, 1991). Therefore, ubiquitin is translated as precursor polypeptides, which need to be processed to produce the free ubiquitin pool for conjugation with other proteins. The resulting polypeptide folds into a barrel like structure, characterised by four antiparallel β-sheets on one side and an α-helix on the other side, and the polypeptide terminates with a flexible C-terminal tail (Figure 1.11). Ubiquitin

contains seven highly conserved lysine (K) residues (K6, K11, K27, K29, K33, K48 and K63). These lysine residues and the N-terminus glycine can all be used for covalent conjugation with other proteins.

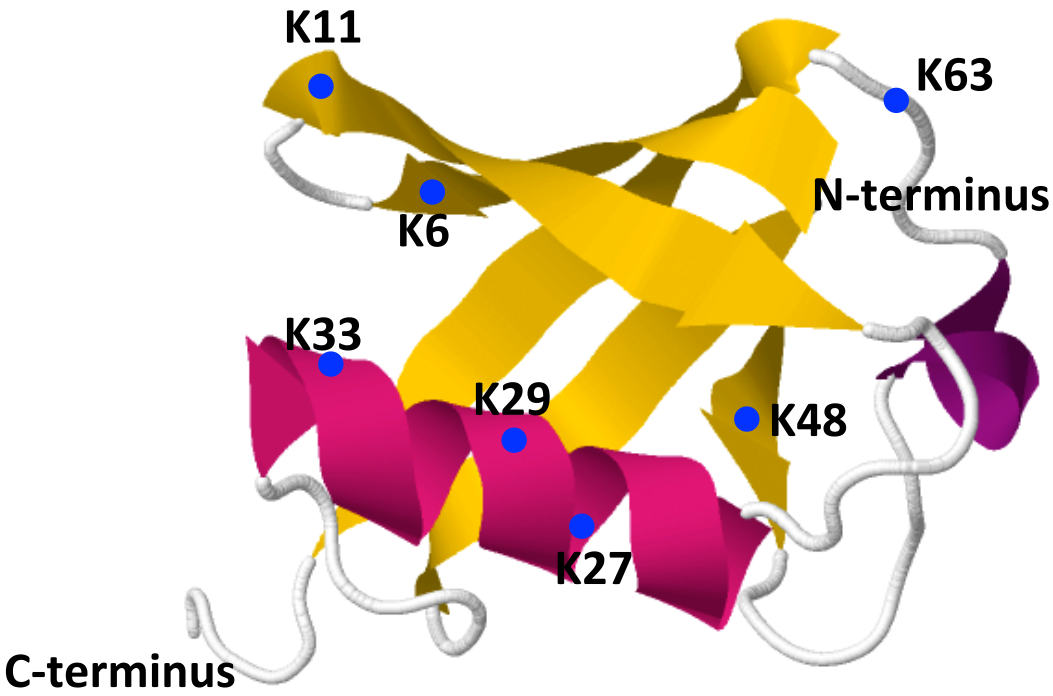

Figure 1.11. Crystal structure of ubiquitin. Ubiquitin is a 76 aa protein comprised of a globular domain containing four stranded mixed β-sheet (in yellow), an α-helix (in pink) and a flexible C-terminal tail terminating in a glycine. Ubiquitin has seven highly conserved lysine residues (blue dots). The structure was retrieved from Research Collaboratory for Structural Bioinformatics Protein Data Bank (PBD ID = 1UBQ).

The biochemical reaction by which ubiquitin is covalently conjugated to another protein is called ubiquitylation. Essentially, it involves formation of an isopeptide bond between the C-terminal glycine residue of ubiquitin and the ε-amino group of a lysine residue of substrate protein. Modification of substrates by ubiquitylation can take many forms (Komander, 2009). Conjugation of a single ubiquitin moiety is termed monoubiquitylation. Accordingly, modification of more than one lysine by a single ubiquitin moiety is referred to as multi-monoubiquitylation whereas modification of substrate proteins by chains of ubiquitin moieties is known as polyubiquitylation. Like other post-translational

modifications (PTM) such as phosphorylation and acetylation, ubiquitylation is a reversible process, where the ubiquitin moieties can be removed by the deubiquitylases (DUBs).

1.2.3.1 The ubiquitylation cascade and ubiquitylating enzymes

Soon after discovery of ubiquitin, it was already demonstrated that ubiquitylation involves sequential actions of three classes of enzymes, namely E1, E2 and E3, which is considered the paradigm of ubiquitylation mechanism (Ciechanover *et al.*, 1982; Hershko *et al.*, 1983).

E1, the ubiquitin-activating enzyme, initiates the cascade of reaction by activating a ubiquitin in an ATP- and Mg^{2+}-dependent manner, where the enzyme binds sequentially to ATP•Mg^{2+} and ubiquitin, and catalyses acyl-adenylation of the C-terminal glycine (Figure 1.12) (Schulman and Harper, 2009). Then, the catalytic cysteine residue of the E1 attacks the ubiquitin-adenylate intermediate, resulting in the formation of a high-energy thiolester bond between itself and ubiquitin. Next, the E1-activated ubiquitin is transferred to the active site cysteine of an E2 conjugating enzyme (Wenzel *et al.*, 2011). In the final step, the ubiquitin is conjugated to the ε-amino group of a lysine residue on the protein substrate in the presence of an E3 ligase. In the case where a HECT (homologous to E6AP carboxy terminus) E3 ligase is engaged, an intermediate thiolester bond is formed between the enzyme and the ubiquitin before the subsequent transfer to the substrate protein. For the case of RING (really interesting new gene) E3 ligase, the E3 enzyme acts as a scaffold to bring the substrate and the E2 in close proximity so that the ubiquitin moiety is transferred directly from the E2 to the substrate. For polyubiquitylation or multi-monoubiquitylation, the discharged E2 is displaced by another charged E2, enabling the transfer of another ubiquitin moiety to the substrate or the substrate-conjugated ubiquitin.

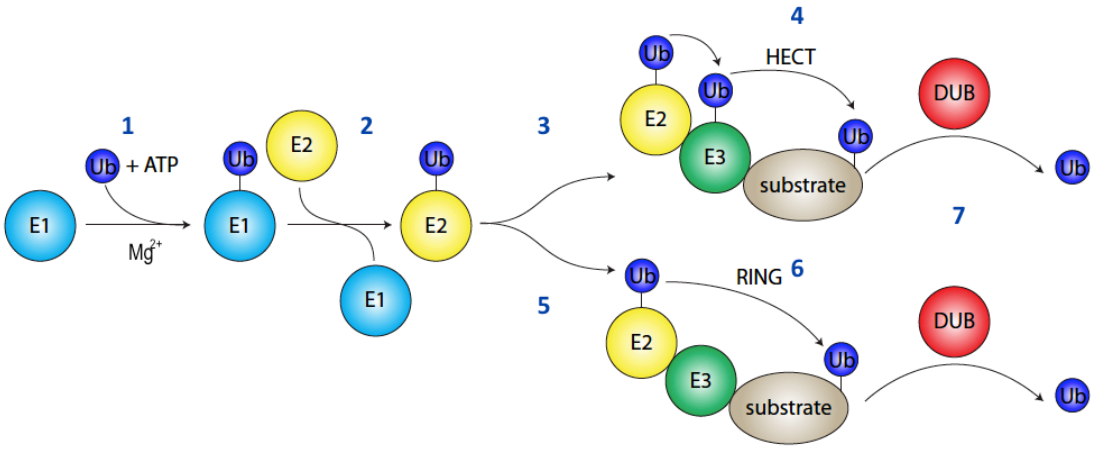

Figure 1.12. Overview of the biochemical steps in protein ubiquitylation and deubiquitylation. (1) Ubiquitin activation by a ubiquitin activating enzyme (E1). **(Ott *et al.*, 1998)** Transfer of ubiquitin to a ubiquitin-conjugating enzyme (E2). **(3)** Transfer of ubiquitin to a HECT domain ubiquitin ligase (HECT E3). **(4)** Association of HECT E3-ubiquitin intermediate with a substrate protein. **(5)** Association of a RING finger ubiquitin ligase (RING E3) with a ubiquitin-carrying E2 and a substrate protein. **(6)** Transfer of ubiquitin to the substrate. **(7)** Deconjugation of ubiquitin from the substrate.

In human, there are two E1s for ubiquitylation, namely UBA1 and UBA6. Structural analysis of the E1 reveals three structural features: a ubiquitin adenylation domain, a Cys domain that contains the cysteine required for thiolester transfer and a ubiquitin-fold domain (UFD) (Schulman and Harper, 2009). The E1 enzymes are highly conserved proteins with prokaryotic antecedents (Figure 1.13). The bacterial proteins molybdopterin converting factor subunit 1 (MoaD) and thiamine biosynthesis protein S (ThiS), which have a ubiquitin-like (UBL) fold (Lake *et al.*, 2001; Wang *et al.*, 2001), are also activated by C-terminal acyl-adenylation by the enzymes molybdopterin biosynthetic enzyme B (MoeB) and ThiF, respectively (Taylor *et al.*, 1998; Leimkuhler *et al.*, 2001), in a similar fashion to ubiquitin adenylation. The human E1s share sequence homology with these bacterial enzymes in the domains responsible for UBL binding and adenylation.

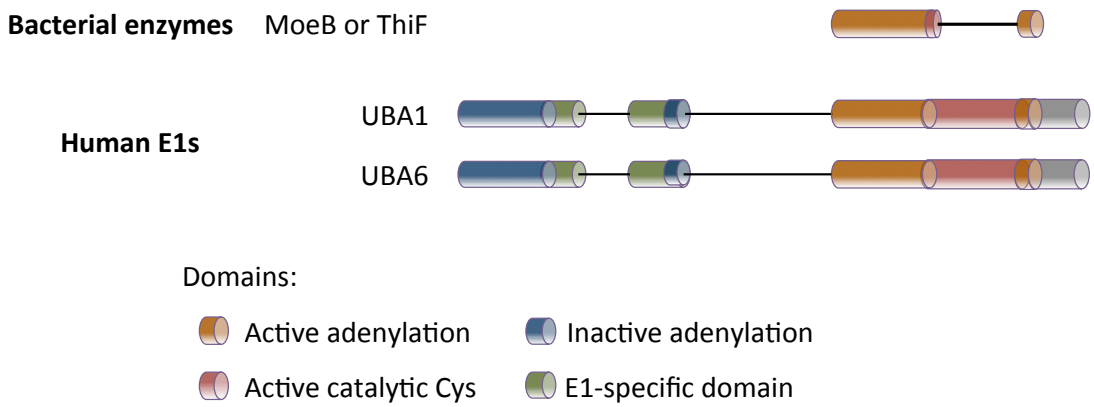

Figure 1.13. Structural homology between the bacterial MoeB or Thif and human E1. The bacterial enzymes molybdopterin biosynthetic enzyme B (MoeB) and ThiF, are characterised by their adenylation domain. The human E1s share sequence homology with these bacterial enzymes in the domains responsible for adenylation, but the human adenylation domains are pseudosymmetrical: one MoeB/ThiF repeat binds to ATP•Mg^{2+} and UBL (active adenylation domain) while the other provides structural stability (inactive adenylation domain). The human E1 also contains two additional E1-specific domains, which are responsible for UBL transfer to E2.

There are at least 38 different E2 enzymes encoded by the human genome (Ye and Rape, 2009; Wenzel *et al.*, 2011). The E2s are characterised by their evolutionarily conserved ubiquitin-conjugating (Ubc) domain, which is 150 amino acids in length. The Ubc domain contains the catalytic cysteine residue for thiolester bond formation and allows interaction with E1s. As initial efforts in the ubiquitin field were focused on E3s, which confer substrate specificity, E2s were long regarded simply as "ubiquitin carriers" with auxiliary roles. However, accumulating evidence indicate their multiple decisive roles in the process of ubiquitylation, including determination of length and linkage-type of ubiquitin chain and regulation of chain assembly reaction. For example, the E2 enzyme UBE2S specifically catalyses formation of Lys11-linked ubiquitin chains (Baboshina and Haas, 1996). Besides, E2s specifically involved in ubiquitin chain elongation often lacks chain-initiating capability and therefore often work in concert to catalyse ubiquitylation. For instance, modification of proliferating cell nuclear antigen (PCNA) during post-replicative DNA (deoxyribonucleic acid) repair with Lys63-linked ubiquitin chain by the E2 Ubc13-Mms2 in yeast (UBE2N-UBE2V1 in human) needs to be preceded by attachment of the first ubiquitin catalysed by the E2 Rad6 (Hoege *et al.*, 2002).

The E3s represent the largest family among the ubiquitylating enzymes and the human genome encodes more than 600 potential E3 ligases (Deshaies and Joazeiro, 2009; Rotin and Kumar, 2009; Metzger *et al.*, 2012). Such a large number of members is necessary to enable specificity in ubiquitin-dependent regulations. There are two major types of E3s in eukaryotes as defined by their conserved catalytic domain, namely the HECT E3 ligases which consist of 28 members (Rotin and Kumar, 2009) and the RING E3 ligases comprising of about 616 members (Deshaies and Joazeiro, 2009). HECT is a C-terminal domain of about 350 amino acids and structural analysis reveals that the HECT has two lobes. The C-terminus lobe of the domain contains the catalytic cysteine residue whereas the N-terminus lobe is responsible for binding to E2s. Region outside of the HECT domain is responsible for substrate binding (Rotin and Kumar, 2009). The RING domain is characterised by a defined pattern of cysteine and histidine residues, which enable binding of two zinc atoms in a cross-brace arrangement (Borden 2000). This zinc ligation is essential for proper folding of the domain, hence for its function. Structure of the RING domain indicates that the E3s are incapable of chemical catalysis but rather function as scaffolding proteins to enable ubiquitin transfer from E2s to substrates. The RING E3s can be subdivided further into those that function as a single subunit and those that act in a multimeric complex, the Cullin RING ligases (CRL) (Li *et al.*, 2008). The proto-oncogene Cbl is a well-studied example of the former and is characterised by two conserved regions: the tyrosine kinase binding (TKB) domain on N-terminus which allows substrate recognition and the C-terminal RING finger domain (Zhang *et al.*, 2000). Cbl is well known for its key role in mediating receptor tyrosine kinases (RTK) ubiquitylation and subsequent degradation (Fukazawa *et al.*, 1996; Miyake *et al.*, 1998). The SCF complex consists of the subunits Skp1, Cullin1 and the F-box protein is a well-known example of the CRLs. In the SCF complex, the Cullin is responsible for E2 recruitment, Skp1 functions as a scaffold while the F-box protein confers substrate specificity (Zheng *et al.*, 2002). There are approximately 70 F-box proteins in human and together they provide recognition of diverse substrates.

1.2.3.2 Different ubiquitin linkages and their physiological roles

As mentioned earlier, all the seven lysine residues and the N-terminus methionine residue in a ubiquitin molecule can be conjugated to the C-terminal glycine of another ubiquitin molecule for extension of a polyubiquitin chain. The different forms of polyubiquitylation (including multiple mono-ubiquitylation and ubiquitin chains linked by different lysine residues) appended to a protein constitutes another layer of control in the regulation of cellular function (Wong and Cuervo, 2010). The differently linked ubiquitin chains adopt different conformation and topologies (Ikeda and Dikic, 2008), and therefore provide different binding surfaces for different ubiquitin-binding proteins (Hicke *et al.*, 2005; Raasi *et al.*, 2005). Accordingly, the attachment of a different ubiquitin linkage dictates a distinct cellular fate of the protein modified.

By virtue of mass spectrometry, it was shown in yeast cells that all the seven lysine residues are used for ubiquitin chain formation, but the different linkages vary in abundance (Peng *et al.*, 2003; Xu *et al.*, 2009). According to data by Xu and colleagues, Lys48- and Lys11-linked ubiquitin chains are the most abundant, followed by Lys63-, Lys6-, Lys27-, Lys29- and Lys33-linked chains (Xu *et al.*, 2009). Besides, they also showed that all the ubiquitin chains, except for Lys63-linked chains can serve as proteolytic signals.

The specific roles of most ubiquitin linkages are still poorly understood. Figure 1.14 summarises the known functions of different ubiquitin chain linkages. The Lys48-linked chain is the best characterised so far as a proteolytic signal (Chau *et al.*, 1989). The Lys63-linked ubiquitin chains are implicated in modulation of cell signalling pathways (such as NFκB (Skaug *et al.*, 2009) and Wnt signalling (Tauriello *et al.*, 2010; Tran *et al.*, 2008)), DNA repair (Stelter and Ulrich, 2003) and receptor endocytosis (Duncan *et al.*, 2006). Monoubiquitylation of substrate has been shown to regulate receptor endocytosis (Hurley and Stenmark, 2011), DNA repair (Huang *et al.*, 2006) and viral budding (Ott *et al.*, 1998). Multiple monoubiquitylation of

receptor tyrosine kinases can also lead to their endocytosis and degradation (Haglund *et al.*, 2003).

Among the unconventional polyubiquitin chains, Lys11-linked chain has been shown to be the preferred linkage type for degradation of cell cycle related proteins, which are the substrate of anaphase promoting complex (APC/C) (Jin *et al.*, 2008; Matsumoto *et al.*, 2010; Wu *et al.*, 2010). Lys11-linked chain is also implicated in endoplasmic reticulum-associated degradation (ERAD) pathway by targeting misfolded proteins for proteasomal degradation (Xu *et al.*, 2009). Lys6-linked chain is implicated in DNA damage and replication stress (Morris and Solomon, 2004). It was shown that upon induction of DNA damage by hydroxyurea, the sites of DNA damage repair is concentrated with proteins modified by Lys6-linked ubiquitin chain and the ubiquitylation relies on the activity of BRCA1.

AIP4, a HECT E3 ligases modifies Deltex, a RING E3 ligase with Lys29-linked ubiquitin chain and targets it for lysosomal degradation (Chastagner *et al.*, 2006). On the other hand, Lys27-linked ubiquitin (together with Lys63-linked chain) is implicated in mitophagy, where the formation of the ubiquitin chain on mitochondrial proteins such as VDAC1 (voltage dependent anion channel 1), is catalysed by the E3 ligase Parkin (Geisler *et al.*, 2010).

Both Lys33-linked and linear ubiquitin chains have been shown to play roles in modulating cell signalling events. The AMPK kinases have been shown to be modified by Lys-33 linked chain (as well as Lys29-linked chain) and such modification leads to inhibition of their activities and phosphorylation by LKB1 (Al-Hakim *et al.*, 2008). These chains can be removed by the DUB USP9X. Linear ubiquitin chain has also been implicated in regulation of NFκB signalling (Haas, *et al.*, 2009; Rahighi *et al.*, 2009; Ikeda *et al.*, 2011; Tokunaga *et al.*, 2011). Modification of NEMO (NFκB essential modulator) by linear ubiquitin chain enhances its interaction with the Tumour Necrosis Factor-R1 signaling complex (TNF-RSC), thereby ensuring efficient TNF-induced activation of NFκB (Haas *et al.*, 2009).

Notably, existence of heterologous ubiquitin chains has also been reported (Ben-Saadon *et al.*, 2006). The Ring1B E3 ligase catalyses its own autoubiquitylation and generates mixture of Lys6-, Lys27 and Lys48-linked ubiquitin chain all on a single ubiquitin molecule. The presence of this mixture of ubiquitylation is required to enable Ring1B to monoubiquitylate histone H2A. Based on these different studies, other than the proteolytic signals encoded by the non-Lys63-linked ubiquitin linkages, all the ubiquitin linkages (except for Lys48-linkages) can lead to diverse cellular outcomes.

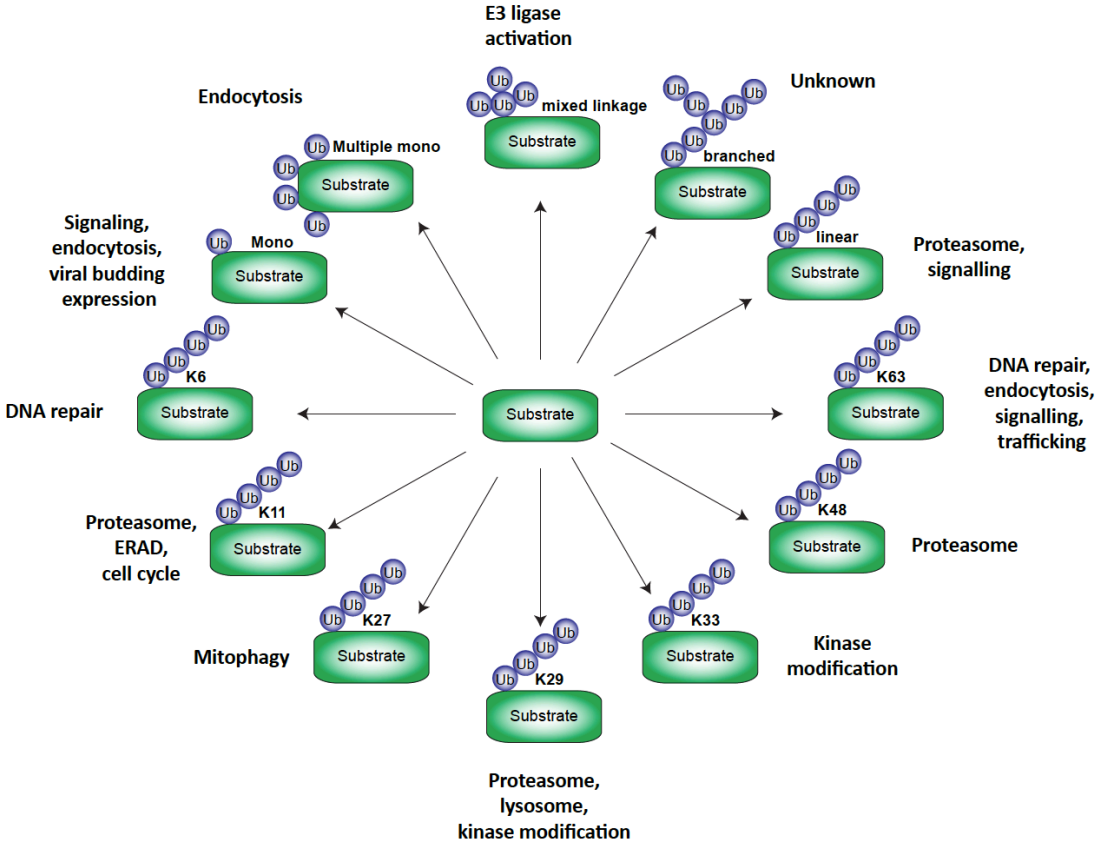

Figure 1.14. Cellular roles of different ubiquitin linkages. All the lysine residues and the N-terminal methionine of ubiquitin can be used for conjugation to form polyubiquitin chains. The different ubiquitin linkages dictate distinct cellular fates of the proteins modified.

1.2.3.3 Ubiquitin binding domains

The use of ubiquitylation to encode diverse signaling functions is likened to the use of Quipus, the talking knots, by people of the Inca Empire as a system of record keeping, in which information was represented by knots along a piece of thread (Komander and Rape, 2012). Just as how the knots can only be deciphered by Quipucamayocs who can read Quipus, the ubiquitin modifications of a substrate require recognition by other proteins to enable the signal to be conveyed. Such a role was fulfilled by the ubiquitin binding domains (UBD) containing proteins, which are capable of recognising the ubiquitylation status of a protein through non-covalent interactions.

The UBDs are found in an expanding number of cellular proteins of diverse roles, which is currently estimated at about 150 (Dikic *et al.*, 2009). To date, more than 20 UBD families are identified and they all differ in term of structure and binding modes (Kirkin and Dikic, 2007). There are UBDs which bind monoubiquitin whereas the others recognize ubiquitin chains of specific linkage type. The UBDs can be broadly categorised into three major classes as dictated by their structural features: α-helix, zinc-finger and β-sheets (Hurley *et al.*, 2006; Kirkin and Dikic, 2007). The α-helix (e.g. UIM and UBA) and zinc-finger UBDs (e.g. NZF and UBZ) are the more common classes of UBDs, whereas the β-sheets UBDs constitute the ubiquitin binding sites in DUBs, E2s and E3s.

By bioinformatic analysis of region of homology among some of the ubiquitylation machineries, the E2, E3 and DUBs, Hofmann and Buher identified the first UBD, the ubiquitin associated domain (UBA) (Hofmann and Bucher, 1996). The UBA has a size of about 55 aa and consists of a compact three-helix bundle (Mueller and Feigon, 2002). A few years later, Wilkinson and colleagues showed that the UBA domain is capable of binding directly to polyubiquitin chains (Wilkinson *et al.*, 2001). On the other hand, the yeast proteasomal subunit S5a was the first ubiquitin-binding protein that is identified biochemically (Young *et al.*, 1998). This discovery offers an explanation for the ability of the 26S proteasome to recognize polyubiquitylated

substrate. The UBD of the yeast S5a was named ubiquitin interacting motif (UIM) and its discovery has enabled identifications of other UIM-containing proteins, including Eps15, Epsin and HRS (Hofmann and Falquet, 2001). Interestingly, these proteins are involved in receptor endocytosis and this early finding therefore highlights a more general role of ubiquitin modification in regulation of protein trafficking. Indeed, these proteins play important roles in recognition of ubiquitylated cargo and their subsequent sorting into the MVB (Williams and Urbé, 2007). A crystal structure of UIM shows that the domain consists of an α-helix centred around an alanine residue (Swanson *et al.*, 2003).

The UBDs differ in their mode of binding to ubiquitin. For instance, binding to monoubiquitin involves interaction with the hydrophobic patch (HP) in the β-sheet of ubiquitin (Dikic *et al.*, 2009). While the UIM and ubiquitin-binding zinc finger (UBZ) domain interact with the HP with a single α-helix apposed next to the central β-sheets, the UBA and coupling of ubiquitin conjugation to ER degradation domain (CUE) bind ubiquitin using two discontinuous α-helices. It is intriguing that an array of helical structures and binding strategies have evolved simply for interaction with ubiquitin. On the other hand, binding of polyubiquitin chain is enabled by positioning UBDs in tandem. Linkage specificity can also be achieved, provided if the tandem UBDs are of the same type, since their avidity are exactly the same so that simultaneous high-affinity interaction is favoured for one linkage type but not the other. For example, BRCA1 which mediates DNA repair has a double UIM domain and shows specificity for Lys63-linked ubiquitin chains (Sobhian *et al.*, 2007; Wang *et al.*, 2007).

1.2.3.4 Deubiquitylases (DUBs)

Another important feature of ubiquitin-dependent regulatory mechanisms is the plasticity and diversity enabled by the DUBs, which are capable of ubiquitin-chain editing or removal. The human genome encodes about 90 DUBs and they fall into 5 categories, namely the ubiquitin specific proteases (USPs), ubiquitin C-

terminal hydrolases (UCHs), Machado Joseph Disease proteins (MJDs), ovarian tumour proteases (OTUs) and JAB1/MPN/Mov4 metalloenzymes (JAMMs) (Komander *et al.*, 2009). Among these, the USPs, UCHs, MJDs and OTUs are cysteine proteases while JAMMs are metalloproteases.

There are three major functional roles of DUBs (Figure 1.15): (1) processing of newly translated polyubiquitin chains or ubiquitin fused to ribosomal proteins to produce free ubiquitin; (2) removal of ubiquitin chains from ubiquitylated proteins, leading to alteration of stability, by rescue from proteasomal or lysosomal degradation, or activity of target proteins. In case where target proteins are committed for degradation, the retrieval of ubiquitin molecules contribute to ubiquitin homeostasis; and (3) editing of ubiquitin modifications by trimming ubiquitin chains (Komander *et al.*, 2009).

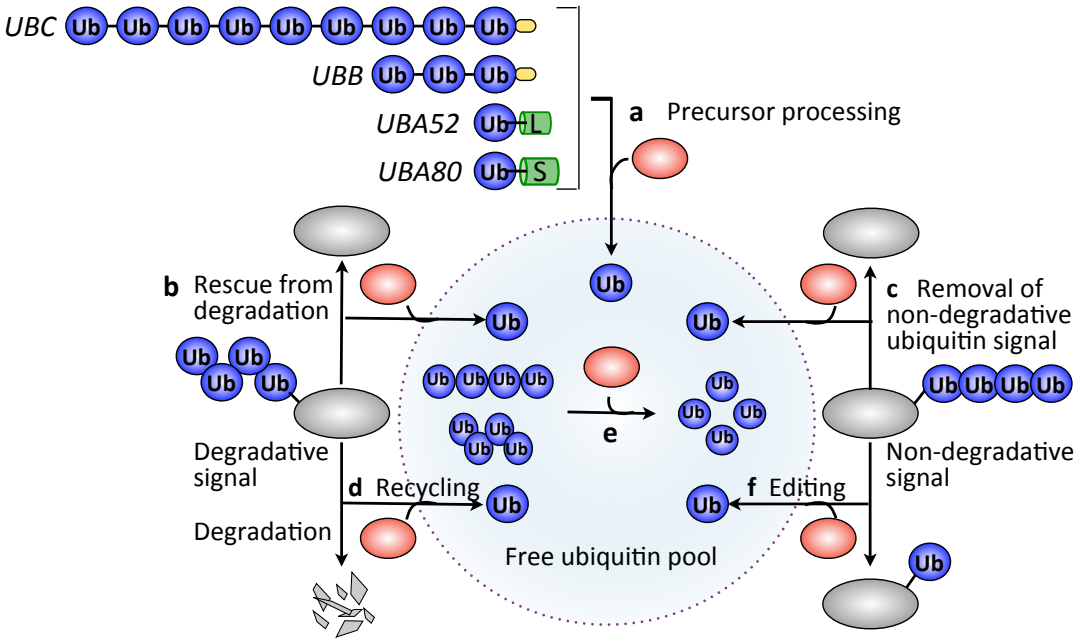

Figure 1.15. Functions of deubiquitylases. Functions of DUBs include **(a)** processing of newly translated polyubiquitin chains or ubiquitin fused to ribosomal proteins to produce free ubiquitin, **(b)** rescuing protein from degradation, **(c)** removing non-degradative signal from substrates, **(d)** maintaining ubiquitin homeostasis through recycling of ubiquitin, **(e)** disassembly of ubiquitin chains generated by en bloc removal from substrates and **(f)** editing ubiquitin chains to exchange one type of ubiquitin signal for the other. (Adapted from Komander *et al.*, 2009)

1.2.3.4.1 Mechanism of deubiquitylation

The catalytic residues of cysteine protease DUBs either consist of a complete catalytic triad comprising a cysteine, a histidine and an aspartic acid or asparagine, or a catalytic diad without the aspartic or asparagine residues (Stoner and Menard 1994; Komander *et al.*, 2009). This catalytic mechanism is highly conserved among the cysteine and serine proteases. In a folded catalytic site, the histidine, which is in close proximity to the cysteine, deprotonates the cysteine residue, enabling its nucleophilic attack on the isopeptide bond between ubiquitin and the conjugated substrate (Figure 1.16). If the aspartic acid or asparagine is present, it aligns and polarizes the histidine. The nucleophilic attack by the thiol group of the cysteine then leads to formation of a negatively charged acyl-intermediate, arising from covalent conjugation of the thiol group with carboxyl group of the conjugated ubiquitin, referred to as the distal ubiquitin. At the final step of the reaction, the acyl-intermediate is hydrolysed by a water molecule, thereby releasing the ubiquitin and the reaction is completed.

The JAMM family DUBs are zinc-dependent metalloproteases, characterised by three highly conserved residues in their catalytic site, namely histidine, asparagine and serine, which coordinate two zinc ions. During catalysis, the zinc ion incorporated in the catalytic site activates a water molecule to form a hydroxide ion, which attacks the carboxyl carbon of the isopeptide bond, leading to formation of a non-covalent intermediate with the substrate. Finally, proton transfer from a nearby water molecule further breaks down the intermediate, resulting in the release of ubiquitin and the DUB (Komander, 2010).

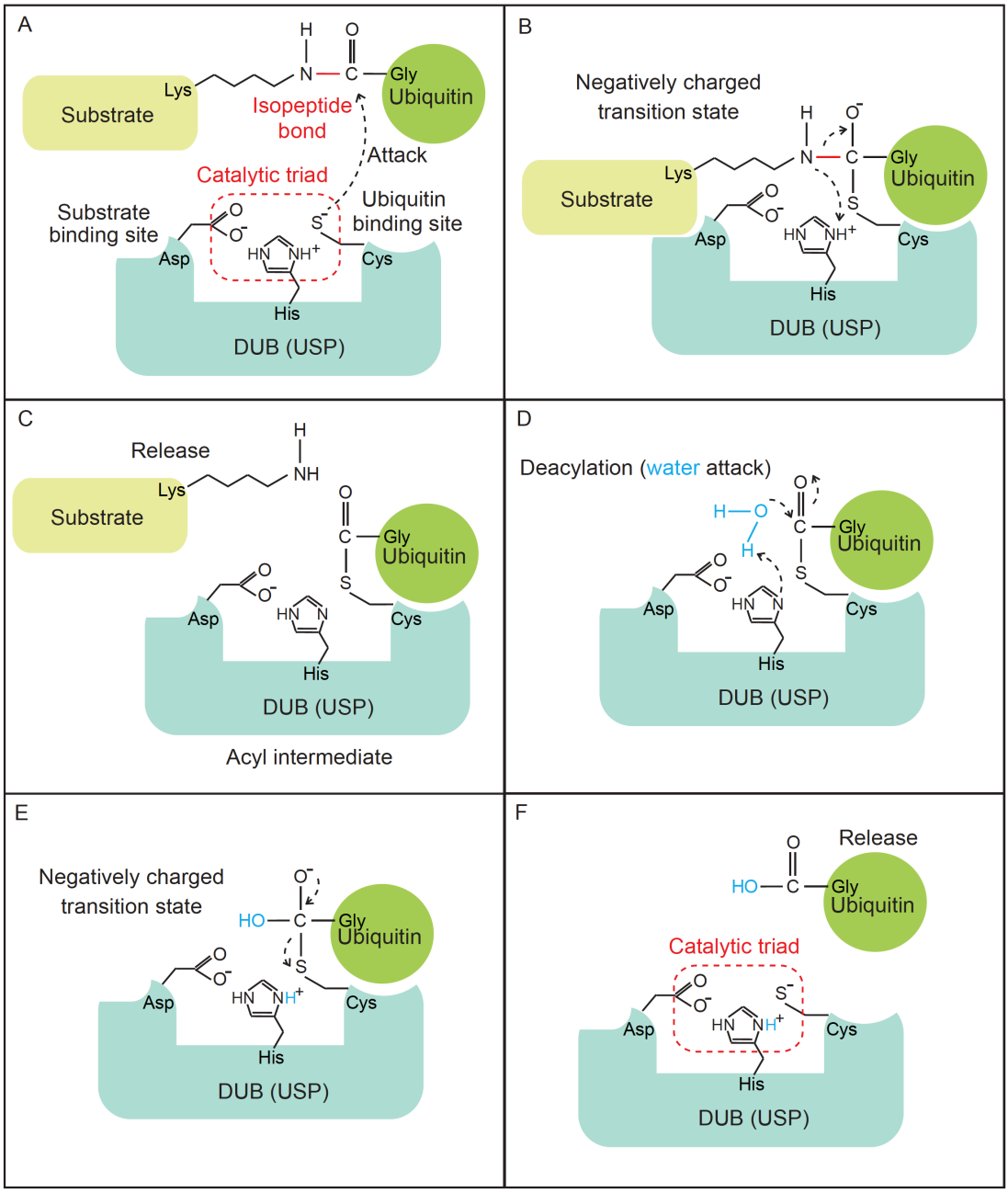

Figure 1.16. Mechanism of deubiquitylation by cysteine protease DUBs. (A) Nucleophilic attack on the carbonyl of the isopeptide bond between the substrate and ubiquitin by the catalytic cysteine. **(B)** The resulting oxyanion intermediate turns into an acyl intermediate, **(C)** while the substrate is being released. **(D)** and **(E)** The hydrolysis of the ubiquitin-DUB intermediate that leads to **(F)** the release of the enzyme and ubiquitin. (Image is a courtesy of Han Liu)

1.2.3.4.2 Substrate specificity of DUBs

The DUBs exhibit specificity for their substrates at multiple levels, in term of their ability to distinguish between ubiquitin and ubiquitin-like molecules, sub-cellular localisation, specific binding interactions and the preference for particular type of ubiquitin chain linkages.

The human genome encodes 17 ubiquitin-like modifiers, which all take a characteristic ubiquitin fold, and ubiquitin is one of them (Reviewed in Hochstrasser, 2009). Among these, SUMO (small ubiquitin-like modifier), Nedd8 (neural precursor cell expressed, developmentally down-regulated 8), ISG15 (Interferon stimulated gene 15) and Atg12 (autophagy related protein 12) are conjugated to their substrates in a similar mechanism to ubiquitin, resulting in topologically similar modifications (Dye and Schulman, 2007; Komander, 2010). However, the DUBs are able to distinguish between these modifications and the selectivity for ubiquitin modifications lies partly in the four residues preceding the diglycine motif at the C-terminus of ubiquitin. SUMO and Atg12 do not share sequence similarity with ubiquitin in this region and are therefore not recognised by DUBs. However, Nedd8 and ISG15 have identical sequence as ubiquitin in this region, so it is not unexpected that both Nedd8 and ISG15 modifications can be hydrolysed by some cross-reactive DUBs, such as USP21 which can process Nedd8 modifications (Gong *et al.*, 2000) and USP2, USP5, USP13, USP14 and USP18, which are crossreactive for ISG15 (Catic *et al.*, 2007).

Many DUBs contain additional protein interaction domains to allow specific sub-cellular localisation or substrate recognition (Komander, 2000). For example, the ubiquitin specific proteases (USPs), USP19 and USP30 have predicted transmembrane domain, which enables them to specifically localise to the ER (Hassink *et al.*, 2009) and mitochondria (Nakamura and Hirose, 2008) respectively. USP19 is implicated in regulation of the unfolded protein response and ER-associated degradation (ERAD) (Hassink *et al.*, 2009) whereas USP30 is involved in regulation of mitochondrial morphology (Nakamura and Hirose, 2008). In another

instance, many DUBs contain additional ubiquitin binding domains, which regulate their activity and/or specificity, and allow them to recognise ubiquitylated substrates (Komander *et al.*, 2009). The UIMs of USP25 (Meulmeester *et al.*, 2008), OTUD25 (Kagayaki *et al.*, 2007) and ATXN3 (Mao *et al.*, 2005) are essential for the activities of the DUBs to efficiently hydrolyse polyubiquitin chains.

Another distinct feature of DUB specificity is the ability of some DUBs to distinguish between different ubiquitin chain linkages (Komander *et al.*, 2009). Our laboratory provided the first description of linkage specificity exhibited by the DUB AMSH (associated molecule with SH3 domain of STAM) (McCullough *et al.*, 2004; McCullough *et al.*, 2006). AMSH (McCullough *et al.*, 2004) and its close relative, AMSH-like protein (AMSH-LP) (Nakamura *et al.*, 2006) have been demonstrated to be stringently selective for Lys63- linked ubiquitin chains. A structural analysis of AMSH-LP bound to a Lys63-linked diubiquitin chain provided important insight of Lys63 specificity of DUBs (Sato *et al.*, 2008). A Lys63-linked polyubiquitin chain has an extended conformation (Komander *et al.*, 2009b) and is stretched maximally upon binding to AMSH-LP. Notably, the sequence context of the Lys63 residue, Gln62 and Glu64, make specificity-determining contacts to the AMSH-LP protein. Other DUBs which have been shown to exhibit linkage specificity include CYLD, which also cleaves Lys63-linked ubiquitin chains (Komander *et al.*, 2008); Cezanne, which cleaves Lys11-linked ubiquitin chains (Bremm *et al.*, 2010); TRABID, which is capable of cleaving Lys29-linked ubiquitin chains (Virdee *et al.*, 2010).

1.2.3.4.3 Ubiquitin specific proteases (USPs)

The USPs represent the largest and most diverse family of DUBs in human, comprising of 56 distinct members with size ranging from 350 to 3400 amino acids and are defined by presence of catalytic USP domain (Komander, 2010). The USP domain varies in size across the USP members, from 295 aa to 850 aa (Ye *et al.*, 2009). Among these, 27 out of the 56 USP domains are less than 400aa. Structural analysis of the USP domain reveal a structure resembling right hand, consisting of

three structural sub-domains: Palm, Thumb and Fingers. The catalytic site lies at the interface between the Palm and Thumb, while the Fingers domain binds to the distal ubiquitin (Komander, 2010).

Multiple sequence alignment of all the USP domains reveals that the core catalytic domain has a length of about 350 aa and there are 6 conserved boxes of amino acid residues. These conserved boxes are interspersed with insertions of amino acids stretches, some of which fold into functional domains like UBL (USP4, USP6, USP11, USP15, USP19, USP31, USP32 and USP43), B-box (CYLD) and ubiquitin binding domains (ubiquitin associated domain (UBA) in USP5 and ubiquitin interacting motif (UIM) in USP37). The regions outside of the USP domains comprise numerous other (predicted) domains, which confer protein binding ability, such as the CAP-Gly domain, UIM and UBA; and which determines subcellular localization, such as transmembrane domain and CAAX-box in USP32 (Komander *et al.*, 2009; Komander, 2010). As mentioned before, USP19 and USP30, which localise to endoplasmic reticulum and mitochondria respectively, both have a predicted transmembrane domain. The UBL domains are also found outside of the catalytic region and there are at least 18 DUBs containing one or more UBL domain (both inside and outside of USP domain), suggesting a general role of this domain in regulation of the DUB activity, which remains to be understood.

1.2.3.4.4 Ubiquitin C-terminal hydrolase (UCH)

The UCH is a small family of DUBs, comprising of 4 members only, namely UCHL1, UCHL3, UCHL5 and BAP1. The UCH domain has a size of about 210 aa. For UCHL1 and UCHL3, the catalytic domain spans almost the entire length of the protein (Komander *et al.*, 2009). UCHL5 has the UCH domain on the N-terminus, followed by an extension of about 100aa, which is important for its incorporation to the proteasome subunit Rpn13 (Hamazaki *et al.*, 2006; Qiu *et al.*, 2006; Yao *et al.*, 2006). BAP1 is the largest member among the UCHs with a size of 730 aa. It has its

UCH domain on the N-terminus, and two predicted nuclear localisation signals (NLS) at the N-terminus.

In vitro studies to assess the DUB activity of UCHL1 and UCHL3 show that the two proteins are unable to process linear, Lys48 or Lys63-linked tetraubiquitin chain (Komander *et al.*, 2009). An explanation was offered when the structure of UCHL3 was revealed that there is a prominent loop covering its active site, hindering substrate accessibility (Johnston *et al.*, 1997). Upon ubiquitin binding, this loop straddles the C-terminal residues of ubiquitin that resides in the UCHL3 active site (Johnston *et al.*, 1999). Due to the presence of this loop, it is unlikely for the UCHs to accommodate ubiquitin chains or ubiquitylated proteins that are too large to fit into the space allowed by the loop. It is therefore proposed that UCHs can only remove ubiquitin from small peptide conjugates, such as the degradative fragments resulting from proteasomal or lysosomal degradation (Komander *et al.*, 2009). However, there is evidence showing activity of UCH enzymes on unfolded large substrates and, accordingly, can potentially deubiquitylate termini of proteins, which can fit into the covered-active site (Johnston *et al.*, 1999).

1.2.3.4.5 Machado-Joseph disease proteins (MJD)

The MJD family also has four members, namely ATXN3, ATXN3L, JOSD1 and JOSD2, characterised by the presence of catalytic Josephin domain, which was first identified by bioinformatics (Scheel *et al.*, 2003). ATXN3 is the most well-known member of the family and is found mutated in Machado-Joseph disease, a form of neurodegenerative disorder termed spinocerebellar ataxia (Riess *et al.*, 2008). All four members of the MJD family are able to hydrolyse Lys48- and Lys63-linked polyubiquitin chains *in vitro* (Weeks *et al.*, 2011). The Josephin domain of ATXN3 was shown to exhibit deneddylase activity against substrate modified by NEDD8, which is a ubiquitin like molecule (Ferro *et al.*, 2007). Together, these observations indicate that Josephins may be promiscuous DUBs exhibiting low level of substrate specificity.

Nuclear magnetic resonance (NMR) studies of the Josephin domain of ATXN3 reveal the presence of a large helical lever that limits accessibility to the active site (Mao *et al.*, 2005; Nicastro *et al.*, 2005). Nicastro and colleagues showed that ATXN3 has two distinct, contiguous ubiquitin binding sites (Nicastro *et al.*, 2009). By performing a NMR-based docking analysis, it was shown that this helical arm is displaced upon docking of a diubiquitin onto ATXN3, suggesting that ubiquitin binding favours an active conformation of ATXN3. Interestingly, ubiquitylation of the Josephin domain itself activates an active conformation of the enzyme (Todi *et al.*, 2009). This leads to the speculation that ubiquitylation renders the helical lever in an open conformation (Komander, 2010).

1.2.3.4.6 Ovarian tumour proteases (OTU)

There are 16 OTU family DUBs in human that have been described to date, and they are cysteine proteases with an incomplete catalytic triad that lacks the aspartate or asparagine residues (Balakirev *et al.*, 2003). The OTU catalytic core has a size between 150 and 200 aa. Although the OTU domain adopts a very different structure from the USP domain, the catalytic residues of the two classes of DUBs superpose well (Komander and Barford, 2008). However, unlike the USPs, the OTU DUBs are more divergent at both sequence and structural levels, and can be further divided into three subfamilies, namely Otubains, OTU-domain (OTUD) and A20-like OTUs (Komander *et al.*, 2009). The Otubains comprise of only two DUBs, namely OTUB1 and OTUB2, and they are the smallest among the OTUs, with a size of less than 280 aa, whereas the OTUDs represent the largest OTU subfamily. The A20-like OTUs are the largest OTUs and have an extended catalytic core of ~360 aa.

Similar to some USP domains, ubiquitin binding to the distal binding site of the OTU domain of yeast OTU1 induces a disorder-to-order transition (Messick *et al.*, 2008). An unproductive configuration of active site has also been reported for OTUB1 and that conformational change is required for activation of the enzyme (Edelmann *et al.*, 2009). Such an inactive conformation of enzyme at resting state is

not unique to OTUs but have also been described in other DUB families and it is proposed to minimise oxidative stress to the catalytic cysteine residue (Komander, 2010).

Multiple members of the OTU DUBs exhibit high chain linkage specificity. TRABID and DUBA preferentially acts on Lys63-linked chains (Kayagaki *et al.*, 2007; Komander et al 2009b) while OTUB1 is Lys48-specific (Edelmann *et al.*,2009).

1.2.3.4.7 JAB1/MPN/Mov4 metalloenzymes (JAMM)

There are 8 members of the JAMM family DUBs and their size varies from 300 to over 2000 aa, but the conserved JAMM domain is only about 100 aa. Multiple members of the JAMM DUBs are found operating as part of a multimeric protein complex (Komander, 2010). POH1 (also known as PSMD14) is associated to the proteasome and plays a role in recycling of ubiquitin chains, thereby contributing to ubiquitin homeostasis (Finley, 2009). AMSH (associated molecule with the SH3 domain of STAM) and AMSH-LP (AMSH-like protein) are associated to the endosomal sorting complex required for transport (ESCRT) and are involved in trafficking of membrane protein (Williams and Urbé, 2007). CSN5 forms part of the COP9 signalosome and deneddylates the Cullin E3 ligase (Cope *et al.*, 2002). MYSM1 is a component of the histone deubiquitinase complex (Zhu *et al.*, 2007). BRCC36 has been found in the BRCA1 A complex involved in DNA repair processes (Dong *et al.*, 2003) and in the BRCC36 isopeptidase complex for which the cellular function is unknown (Cooper *et al.*, 2009).

Interestingly, six of the eight human JAMMs, (CSN5, AMSH (McCullough *et al.*, 2004), AMSH-LP, BRCC36 (Cooper *et al.*, 2009), POH1 and MYSM1) are able to process Lys63-linked ubiquitin chains (Komander *et al.*, 2009). Given the diverse roles of Lys63 ubiquitin linkages, the linkage specificity of the JAMMs suggests their important regulatory roles.

1.2.3.5 Cellular roles of ubiquitylation

Ubiquitylation is involved in the regulation of diverse cellular processes, such as proteasomal-mediated protein degradation, endocytic trafficking, transcriptional regulation and cell signalling.

1.2.3.5.1 Proteasomal-mediated protein degradation

Ubiquitylation was first identified as a post-translational modification that targets protein for degradation by the proteasome (Ciechanover, 2005). The proteasomal-mediated protein degradation is important for cellular processes such as protein quality control, antigen processing, signal transduction, cell cycle control, cell differentiation and apoptosis (Voges *et al.*, 1999). Besides, the proteasome also serves an important function of retrieving ubiquitin prior to degradation of the polyubiquitylated substrates, thereby contributing to ubiquitin homeostasis (Komander *et al.*, 2009). This is achieved by the activities of the three proteasomal-associated DUBs, namely POH1, USP14 and UCHL5, which can remove ubiquitin chains from the substrates. Besides, there is also a growing notion that the DUBs provide a proof-reading mechanism by facilitating release of substrate proteins from the proteasome if they are not commited to degradation within a given time window (Clague and Urbé, 2010). Jacobson and colleagues demonstrated that the levels of Lys63-linked ubiquitin chains are unaffected by short-time inhibition of proteasomes. While both Lys48- and Lys63-linked ubiquitin chains bind to proteasome with comparable affinities *in vitro*, Lys63-linked ubiquitin chains are deubiquitylated six-fold more rapidly than Lys48-linked ubiquitin chains. Moreover, UbcH10 conjugated to Lys48-linked tetraubiquitin chain is targeted for proteasomal degradation whereas Lys63-linked tetraubiquitin chain conjugated to UbcH10 is rapidly deubiquitylated but UbcH10 is not efficiently degraded. Together, these data suggest that the preferential proteasomal DUB activity towards Lys63-linked ubiquitin chains is selecting against Lys63-chain modified substrates for proteasomal degradation (Jacobson *et al.*, 2009).

1.2.3.5.2 Endocytic trafficking

The first evidence for the role of ubiquitylation in membrane trafficking came from the work of Hicke and Riezman, which demonstrated that efficient endocytosis of Ste2p, a G-protein coupled membrane receptor in yeast, requires ubiquitylation of the protein (Hicke and Riezman, 1996). This was later shown to be a signal for endocytosis of other membrane proteins, such as the yeast uracil permease Fur4 (Galan *et al.*, 1996), growth hormone receptor in Chinese hamster (Strous *et al.*, 1996) and epidermal growth factor receptor in human (Stang *et al.*, 2004). Unlike proteasomal degradation of protein, which requires polyubiquitylation of the substrate, monoubiquitylation suffices as a signal for endocytosis (Terrell *et al.*, 1998; Haglund *et al.*, 2003).

While ubiquitylation was first discovered to be a signal for endocytosis, it was later revealed that it plays a more prominent role as a signal for sorting of membrane proteins into the intraluminal vesicles (ILVs) during formation of multivesicular bodies (Hurley and Stenmark, 2011). Following endocytosis, the ubiquitylated membrane protein trafficks to the early endosome and is captured by the endosomal sorting complex for transport (ESCRT) machineries. The ESCRT machineries then mediate the sorting of the ubiquitylated cargo protein into ILVs. Multiple components of the ESCRT complexes contain ubiquitin-binding domains, which allow them to interact with the ubiquitylated cargo and therefore ensures efficient concentration of the cargo into a MVB (Railborg and Stenmark, 2009). Both hepatocyte growth factor regulated substrate (Hrs) and signal transducing adaptor molecule (Swanson *et al.*,2003) of ESCRT-0 bind ubiquitylated cargo via their UIMs. Tsg101 (tumour suppressor gene 101) of ESCRT-I binds ubiquitin via its ubiquitin E2 variant (UEV) domain whereas ubiquitin binding of Vps36 of ESCRT-II is conferred by its GLUE (GRAM-like ubiquitin-binding in Eap45) domain.

Although ubiquitylation signals sorting of proteins into the MVBs, the deubiquitylation of the cargo by a DUB is also essential for efficient sorting (Raiborg and Stenmark, 2009). In mammalian cells, two DUBs of different families, AMSH and

USP8 (also known as UBPY), are found recruited to ESCRT-0 (through interaction with STAM) (McCullough *et al.*, 2006; (Mizuno *et al.*, 2006; Row *et al.*, 2006), ESCRT-II (through interaction with Alix) (Luhtala and Odorizzi, 2004) and ESCRT-III complexes (through interaction with CHMP3) (Agromayor and Martin-Serrano, 2006). The DUBs may play a role in promoting disengagement of the cargo from the ESCRT machineries or to deubiquitylate cargo that is not destined for sorting into MVB (Williams and Urbé, 2007; Raiborg and Stenmark, 2009). Proteins which are sorted into the MVB are eventually degraded via the lysosomal pathway.

Several receptor tyrosine kinases (RTKs), such as EGFR, Met receptor for HGF and platelet-derived growth factor receptor (PDGFR), upon stimulation by their cognate ligands, become ubiquitylated by the RING E3 ligase Cbl (Marmor and Yarden 2004). The polyubiquitylation results in their endocytosis and subsequent sorting into the MVB before fusion with lysosome for degradation. Therefore, the ubiquitin-mediated endocytosis and sorting, serve as an important mechanism to prevent persistent activation of the signalling pathways.

1.2.3.5.3 Ubiquitin and transcriptional regulation

Both the proteolytic and non-proteolytic function encoded by ubiquitylation are implicated in the process of transcriptional regulation (Geng *et al.* 2012).

The level of transcriptional activators, repressors and their coactivators need to be tightly regulated in order to fine-tune the transcriptional programme that they regulate (Muratani and Tansey, 2003; Geng *et al.*, 2012). This is achieved by the activity of the UPS to degrade these proteins when their functions are not required. One of the best-studied examples of this is the regulation of β-catenin level (see section 1.1.2.2.2 for detailed description). There are also specific examples where the transcriptional activators are subjected to limited proteolysis by the UPS to become functional. For example, the C-terminus of the p105 precursor protein is

degraded by the UPS, leaving the N-terminal p50, which becomes part of the NFκB transcription complex (Palombella *et al.*, 1994).

Non-proteolytic ubiquitylations are also implicated in regulation of transcription (Geng *et al.*, 2012). There are multiple mechanisms by which ubiquitylation can influence activities of transcriptional proteins. Firstly, ubiquitylation can affect localisation of the transcription proteins. For example, the monoubiquitylation of FOXO4, a transcription factor induced upon cellular stress, results in its translocation to the nucleus for transcriptional activation while its deubiquitylation by USP7 results in its nuclear exclusion (van der Horst *et al.*, 2006). Ubiquitylation has also been proposed to allow temporal control of transcriptional activity. SRC-3 (steroid receptor coactivator-3), a coactivator for hormone-liganded transcription factor, becomes activated upon monoubiquitylation (Wu *et al.*, 2007). The extension of the ubiquitin chain on SCR-3 beyond a certain limit renders it for proteasome mediated degradation. The time taken for extension of the polyubiquitin chain prior to degradation therefore acts as a "molecular clock" to define a time window for which SRC-3 can activates transcription. There is also evidence to indicate ubiquitylation affects binding affinity of transcription proteins to their target sites, both positively and negatively (Geng *et al.*, 2012). Monoubiquitylation of the yeast activator Gal4 enables it to bind firmly to chromatin without being stripped off by the ATPases resident in the 19S base complex of proteasome (Archer *et al.*, 2008).

Besides, ubiquitylation also affects chromatin structure and function through its regulatory role on histones. It is evident now that histone ubiquitylation is abundant within cell and is important for control of gene activity (Weake and Workman 2008). Ubiquitylation of histone H2A is typically implicated in compaction of chromatin, thereby repressing transcription whereas H2B ubiquitylation is normally implicated for transcriptional activation (Geng *et al.*, 2012).

1.2.3.5.4 Ubiquitin-dependent regulation of the canonical Wnt signalling pathway

Ubiquitylation is also implicated in regulation of different aspects of cell signalling pathways, by regulating its stability of signalling components, activating downstream signalling through proteolytic processing and providing a surface of interaction through ubiquitin chains (Haglund and Dikic, 2005). While the role of ubiquitylation in cell signalling is best exemplified by its implication in the NFκB signalling pathway (Skaug *et al.*, 2009), there is also an expanding repertoire of literature reporting the ubiquitin-dependent regulation of the Wnt signalling pathway.

Multiple components of the canonical Wnt signalling pathway are subjected to modification by ubiquitylation, which alters their stability, binding affinity, localisation and activity (Tauriello and Maurice, 2010). As mentioned before (see section 1.1.2.2.2), the level of β-catenin, the key signal transducing molecule of canonical Wnt signalling pathway, is tightly regulated by the UPS and its ubiquitylation is mediated by βTrCP (Aberle *et al.*, 1997). In fact, there are also other E3 ligases capable of ubiquitylating β-catenin. Siah-1, a RING E3 ligase, which is upregulated upon genotoxic stress-mediated p53 activation, ubiquitylates β-catenin in a phosphorylation independent manner and targets it for proteasomal degradation (Liu *et al.*, 2001; Matsuzawa and Reed, 2001). Jade-1, is another E3 ligase that ubiquitylates β-catenin in a phosphorylation-dependent manner, such that GSK3β activity is a prerequisite for its action (Chitalia *et al.*, 2008). However, unlike βTrCP-mediated regulation of β-catenin, Jade-1-mediated regulation of β-catenin is involved both in the absence and presence of a Wnt ligand (Tauriello and Maurice, 2010). However, it was recently revealed that, ubiquitylation on β-catenin does not always entail proteolysis. The EDD (E3 ligase identified by differential display), a HECT E3 ligase, ubiquitylates β-catenin, which leads to its enhanced stability and activity, thereby upregulating Wnt signalling (Hay-Koren *et al.*, 2011).

Importantly, the EDD catalyses formation of Lys11- or Lys29-linked polyubiquitin chain, instead of the proteolytic-encoding Lys48-linked chain.

Besides, both APC and Axin, the components of the destruction complex of β-catenin, are regulated by ubiquitylation. Similar to β-catenin, the ubiquitylation of APC by EDD (discovered 4 years before role of EDD on β-catenin stabilisation was known), results in its enhanced stabilisation and facilitates β-catenin downregulation when Wnt ligand is absent (Ohshima *et al.*, 2007). Conversely, upon activation by Wnt ligand, APC is ubiquitylated by a yet-to-be identified E3 ligase and is subsequently degraded by the proteasome (Choi *et al.*, 2004). On the other hand, the COP9 signalosome associated DUB, USP15 has been shown to be the DUB that deubiquitylates APC and leads to its stabilisation (Huang *et al.*, 2009). Recently, Tran and colleagues identified TRABID, an OTU DUB, regulates the Wnt signalling pathway through its activity on APC (Tran *et al.*, 2008). siRNA depletion of TRABID leads to hyperubiquitylation of APC (possibly by Lys63-linked ubiquitin chains given the strong preference of TRABID for this linkage type) but does not affect APC stability (nor that of β-catenin). Of note, APC has been shown to also act downstream of β-catenin stabilisation and repress β-catenin mediated Wnt signalling (Sierra *et al.*, 2006). Therefore, it is suggested that ubiquitylation of APC by Lys63-linked ubiquitin chain leads to enhanced repressive activity mediated by APC. On the other hand, Stability of Axin is regulated by USP34, such that siRNA depletion of USP34 results in downregulation of Axin (Lui *et al.*, 2011). The stabilisation of Axin, in this case, is however required for β-catenin-mediated transcription activation but the mechanism of this warrants further research efforts.

Dishevelled (Dvl), which is important for both canonical and non-canonical Wnt signalling, also undergoes ubiquitin-dependent regulation. The ubiquitylation of Dvl by Kelch-like 12 (KLHL12)-Cullin3 E3 ligase complex following Wnt stimulation results in its proteasomal degradation (Angers *et al.*, 2006). Recently, ITCH, a HECT-containing Nedd4-like ubiquitin ligase, has also been shown to ubiquitylate Dvl in a phosphorylation dependent manner and renders it for proteasomal degradation

(Wei *et al.*, 2012). In both cases, the ubiquitylation and degradation of Dvl represent a negative feedback mechanism to downregulate Wnt signalling following activation of the pathway. In addition to regulating its stability, the ubiquitylation of Dvl by Lys63-linked chain, has been associated with Dvl polymerisation and stabilisation of the complex, and is important for activation of Wnt signalling pathway (Tauriello *et al.*, 2010). The Lys63-linked chain appended to Dvl can be removed by a tumour suppressor DUB, CYLD, which is mutated in cylindroma tumor, thereby activating Wnt signalling.

Recently, two independent groups published in Nature, the identification of two homologous transmembrane RING E3 ligases, namely RNF43 (ring finger 43) and ZNRF3 (zinc and ring finger 3) for their roles in regulating the Wnt receptor complex (Hao *et al.*, 2012; Koo *et al.*, 2012). Both of the E3 ligases are acting as negative regulators of the Wnt signaling pathway, by promoting ubiquitylation and subsequent degradation of the Frizzled receptor and LRP6 co-receptor.

1.2.3.6 Deubiquitylases and cancer

Many proteins regulated by the UPS are involved in cellular processes germane to tumourigenesis and tumour progression, such as cell-cycle progression, cell adhesion, apoptosis, endocytic trafficking of receptors and gene transcription (Hoeller and Dikic, 2009). Among the multiple components of the UPS, the role of DUBs in cancer is of particular interest to our lab. Since ubiquitylation of a protein can affect its stability or signalling function, the removal of the appended ubiquitin(s) therefore also affects either of these two aspects, but in the opposite direction (Sacco *et al.*, 2010).

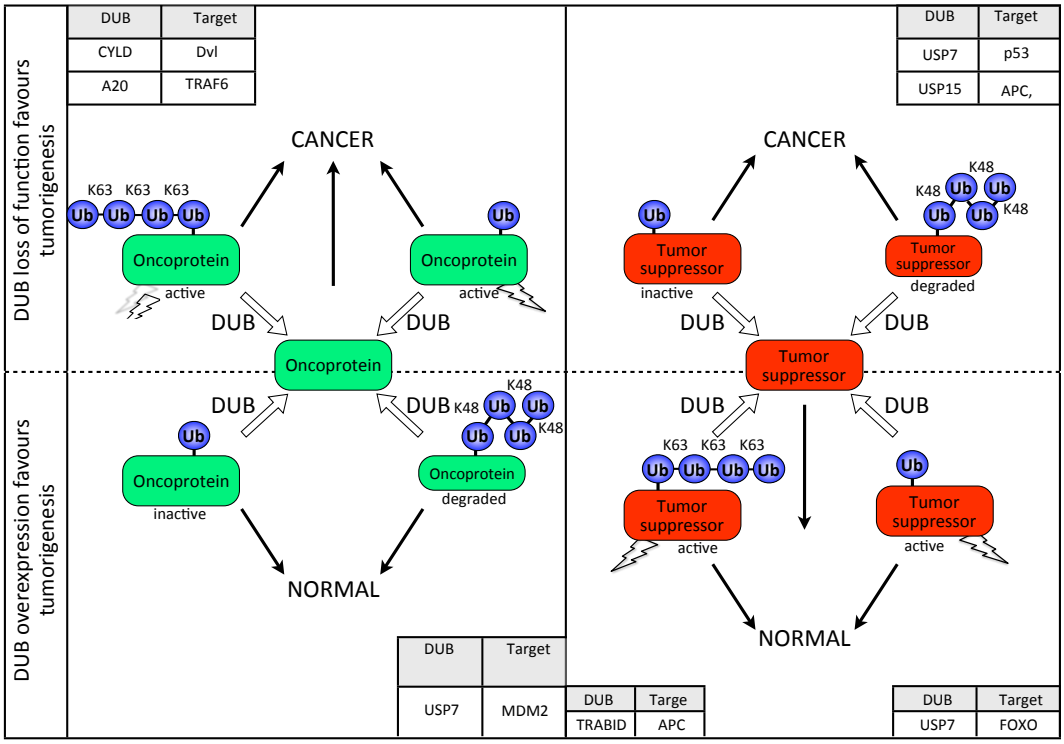

Figure 1.17. Deubiquitylases are important regulators of oncogenes and tumour suppressor proteins. The loss of function of DUBs and their overexpression can both lead to development of cancer. DUBs regulate the stability of an oncoprotein (e.g. Mdm2) or a tumour suppressor by rescuing it from proteasomal degradation (e.g. p53). Also, DUBs can regulate their activity (activation or inactivation) by editing the chains appended to the substrates. Activation may refer to engagement in signalling protein network (e.g. Dvl) or translocation to the nucleus (e.g. FOXO). Examples of DUBs implicated in cancer development are shown with their cognate targets.

Considering the stabilising effect of DUBs by rescuing proteins from degradation, in the context of cancer, the overexpression of a DUB regulating an oncogenic protein therefore favours tumour progression and *vice versa*, while the opposite is true for a tumour suppressor protein (Figure 1.17). The regulation of the tumour suppressor p53 and its E3 ligase, MDM2, by USP7 is a good embodiment of this mode of regulatory mechanism. Ubiquitylation of p53 by MDM2 and autoubiquitylation of MDM2 result in their proteasomal degradation. Therefore, USP7 has a dual effect on p53 stability, depending on whether it predominantly deubiquitylates p53 or MDM2. Volgelstein's group demonstrated that under physiological condition, MDM2 is the preferred substrate of USP7 since the loss of the DUB leads to a stabilisation of p53 while MDM2 is degraded (Cummins *et al.*, 2004; Cummins and Vogelstein, 2004). In case where a ubiquitylation signal is

required for the activation of a signalling pathway, such as the case of Lys63-linked modification of Dvl which promotes Wnt signalling upstream of β-catenin, the removal of the ubiquitin chain by CYLD leads to suppression of the pathway (Tauriello *et al.*, 2010). On the other hand, downstream of β-catenin in the Wnt signalling pathway, the ubiquitylation of APC by Lys63-linked chain actually leads to attenuation of Wnt signalling and deubiquitylation of APC by TRABID leads to activation of the pathway (Sierra *et al.*, 2006).

1.2.3.6.1 Emerging roles of BAP1 in cancer

BAP1 is a DUB, the mutations of which, has been associated with a wide range of cancers, revealing its tumor suppressive role (White and Harper, 2012). It also turns out to play a regulatory role on β-catenin, a proto-oncogene based on my work discussed in chapter 4.

BAP1 (BRCA1-associated protein 1) is the largest member of the UCH family of DUB comprising of 729 amino acids. It was first identified as an interacting partner of the tumour suppressor BRCA1 (breast and ovarian cancer susceptibility protein 1) (Jensen *et al.*, 1998; Jensen and Rauscher, 1999). The first 240 aa of BAP1 comprises its UCH domain and there are two putative nuclear localisation signals (NLS) close to its C-terminus. It also has a predicted coiled coil domain at the C-terminus (Figure 1.18).

The role of BAP1 as a tumour suppressor was already recognised at the time when it was discovered (Jensen *et al.*, 1998). Its activity was required for BRCA1-mediated inhibition of breast cancer cell growth and intragenic homozygous rearrangement and deletions of BAP1 was detected in lung carcinoma cell lines. However, BAP1 was only confirmed as a *bona fide* tumor suppressor two years ago following exome sequencing of metastatic uveal melanoma (UM) (Harbour *et al.*, 2010). BAP1 was found mutated in only one of twenty-six samples of class I (low metastatic risk) UM but was found mutated in twenty-six out of thirty-one class 2

UM (high metastatic risk) samples that were analysed. siRNA depletion of BAP1 in 92.1 UM.1 cells, which did not harbour detectable BAP1 mutation, led the cells to acquire features reminiscent of class 2 clinical biopsy samples. Together, these suggest that loss of BAP1 is a prerequisite for acquisition of metastatic state of the cancer. Importantly, in this study, it was revealed that BAP1 mutations are heterogenous and can be detected along the gene (Figure 1.19). It is intriguing how these heterogenous mutations can all lead to the same biological outcome. The study by Harbour and colleagues was followed by detection of BAP1 somatic mutations in malignant pleural mesothelioma (Bott *et al.*, 2011) and renal cell carcinoma (Pena-Llopis *et al.*, 2012), and germline mutations in familial cancer predisposition syndromes such as malignant mesothelioma (Testa *et al.*, 2011) and melanocytic tumours (Wiesner *et al.*, 2011). More recently, using a BAP1-deficient mouse model, Dey and colleagues demonstrated that the loss of BAP1 led to development of myeloproliferative disorder, which is reminiscent of human chronic myelomonocytic leukaemia (CMML) (Dey *et al.*, 2012).

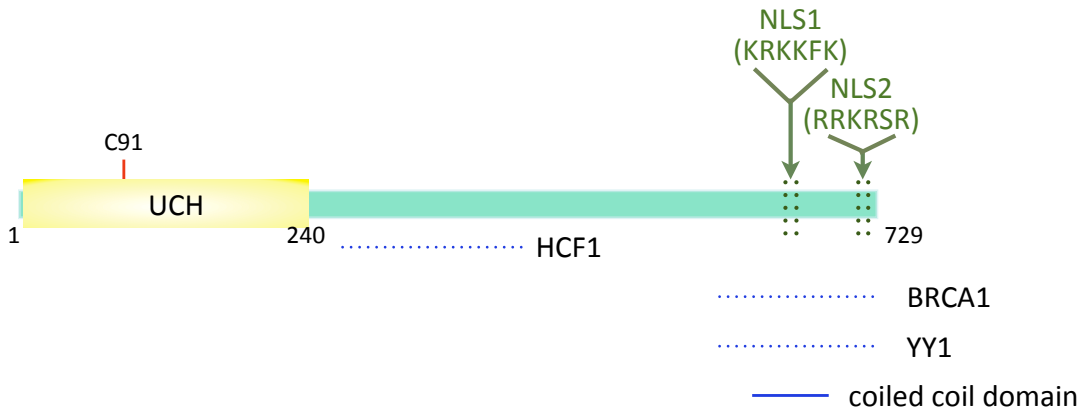

Figure 1.18. Domain structure of BAP1. BAP1 has a conserved UCH domain on its N-terminus and two nuclear localisation signals (NLS) are found at its C-terminus. The C-terminus also has a predicted coiled coil domain. Binding sites for known interacting partners of BAP1, namely HCF1, BRCA1 and YY1, are indicated by dotted line.

The biology of BAP1 in regulation of cellular processes is still emerging. By introducing BAP1 into NCI-H226, a non-small cell lung cancer in which BAP1 is

homozygously deleted, Ventii and colleagues showed that BAP1 exerts growth suppressive effect on the cell line by affecting the cell cycle, such that progression through the G_1-S checkpoint is sped up, and subsequently results in induction of cell death (Ventii *et al.*, 2008). Importantly, the growth suppression mediated by BAP1 is dependent on its enzymatic activity and nuclear localisation. The promotion of G_1/S transition during cell cycle by BAP1 has also been demonstrated in HeLa (Nishikawa *et al.*, 2009) and MCF10A cells (Machida *et al.*, 2009). The role of BAP1 on cell cycle progression is likely mediated through its regulatory role on HCF-1 (Host cell factor 1), which is known to play a role in transcriptional control of cell cycle regulation (Tyagi *et al.*, 2007; Eletr and Wilkinson, 2011). Three independent *in vitro* studies (Machida *et al.*, 2009; Sowa *et al.*, 2009; Yu *et al.*, 2010) and an *in vivo* study (Dey *et al.*, 2012) have all identified HCF-1 as the key interacting partner of BAP1, and other proteins also found in the complex include ASXL1 (additional sex combs like 1), ASXL2, OGT (O linked-N-acetylglucosamine transferase), FOXK1 (forkhead box K1) and FOXK2. BAP1 is acting as part of a transcriptional complex that regulates gene transcription (Yu *et al.*, 2010). Yu and colleagues showed that siRNA depletion of BAP1 leads to deregulation of expression of 1244 genes, which are involved in wide range of cellular processes including cell cycle, DNA replication and repair, cell death, cellular growth and proliferation and cellular assembly. They further demonstrated that BAP1 associates with transcriptionally active chromatin and forms a ternary complex together with HCF-1 and YY1 (Yin Yang 1) to regulate gene expression relevant to cell cycle progression.

These studies provide a foundation, which further understanding on the role of BAP1 in tumour suppression can be built upon. It is intriguing how the wide range of genes affected by BAP1 coordinate and culminate in consistent tumour suppression effect in different cancer types.

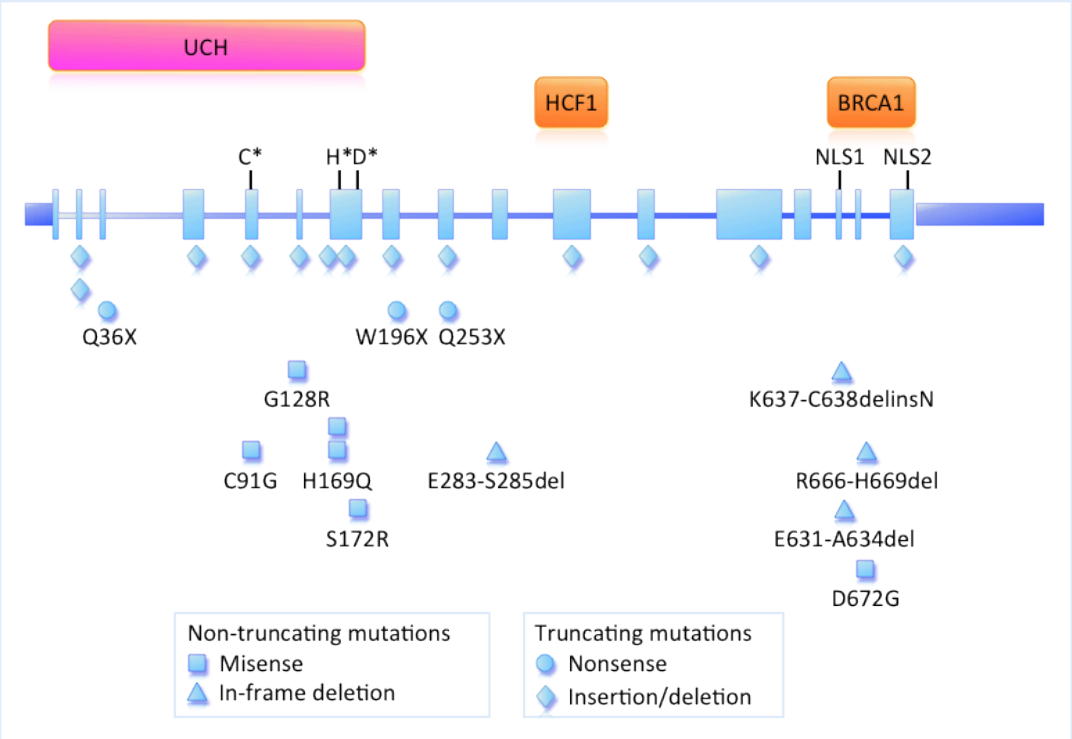

Figure 1.19. Map of BAP1 gene and positions of BAP1 mutations detected in uveal melanoma. BAP1 contains 17 exons (light blue boxes) that encode a 729 aa protein and the mutations detected are shown below the figure as indicated. Region corresponding to UCH domain of BAP1 (pink box) and regions involved in binding to HCF1 and BRCA1 (orange boxes) are shown above the figure as indicated. The critical Cys91 (C), His169 (H) and Asp184 (D) residues in the catalytic sites and nuclear localisation signals (NLS1 and 2) are indicated (Harbour *et al.*, 2010).

1.3 RNA interference

In 1984, Mizuno *et al.* provided the first evidence in *Escherichia coli* that ribonucleic acid (RNA) can regulate gene expression post-transcriptionally via complementary binding of a RNA transcript to the corresponding messenger RNA, thereby inhibiting protein translation. About a decade later, two groups independently identified the gene lin-4, which encodes small RNA (later termed microRNA by Victor Ambros) that hybridises to 3' UTR of lin-14 gene owing to its sequence complementarity and regulates lin-14 expression via a post-transcriptional mechanism during the development of Caenorhabditis elegans (Lee *et al.*, 1993;

Wightman *et al.*, 1993). The significance of this finding is however not well appreciated until a second member of such small RNA, let-7 was identified again in C. elegans and more importantly, similar sequences were found in other species (Reinhart *et al.*, 2000).

In parallel to this, a phenomenon called post-transcriptional gene silencing (PTGS) was first described in 1990, where an exogenously introduced cloned gene into the plant genome was found to inhibit expression of homologous gene sequences and considerable amount of evidence were demonstrated by subsequent investigations of viral infection in plants, providing clues to the mechanisms of PTGS. This was followed by verification of similar process called quelling in the fungus Neurospora crassa.

These independent findings provided clear evidence indicating the roles of RNA in gene regulation but the haze of these gene silencing events were not cleared until the ground breaking discovery of RNA interference (RNAi) by Andrew Fire and Craig Mello in 1998 (Fire *et al.*, 1998). In essence, in their seminal paper published in Nature, Fire and colleagues demonstrated that double stranded RNA (dsRNA), but not single stranded RNA, which correspond to a mature mRNA sequence, is very potent to cause gene silencing by inducing degradation of the endogenous mRNA, in a non-stoichiometric fashion and such effect is transmissible between tissues and progenies.

Ever since then, there was an explosion of information for the mechanisms of RNAi. A hallmark of RNAi is the generation of short (20-30 nucleotides) dsRNA by RNAseIII enzymes (Siomi and Siomi, 2009). In general, small RNAs can be categorized into two classes based on their origin, namely siRNAs generated by cleavage of long exogenous long dsRNA precursors and microRNAs generated following processing of genome-encoded stem-loop structures. The siRNAs generated by Dicer are ~21-25 nucleotide duplexes flanked by a phosphate group on both of its 5' ends, and the 3' ends are flanked by nucleotide. The binding of Dicer to dsRNA is mediated by its PAZ (Piwi Argonaut and Zwille) domain and the processing of dsRNAs into siRNAs is

catalysed by its two RNAseIII domains (Zhang *et al.*, 2004). The siRNA duplex is subsequently loaded onto Argonaut protein and is incorporated into the RISC (RNA-induced silencing complex) complex, where the duplex is unwound, leaving a single stranded guiding siRNA which guides the complex by sequence complementarity to target mRNA and the RISC complex then cleaves the target mRNA, thereby preventing protein expression of the target gene (Figure 1.20).

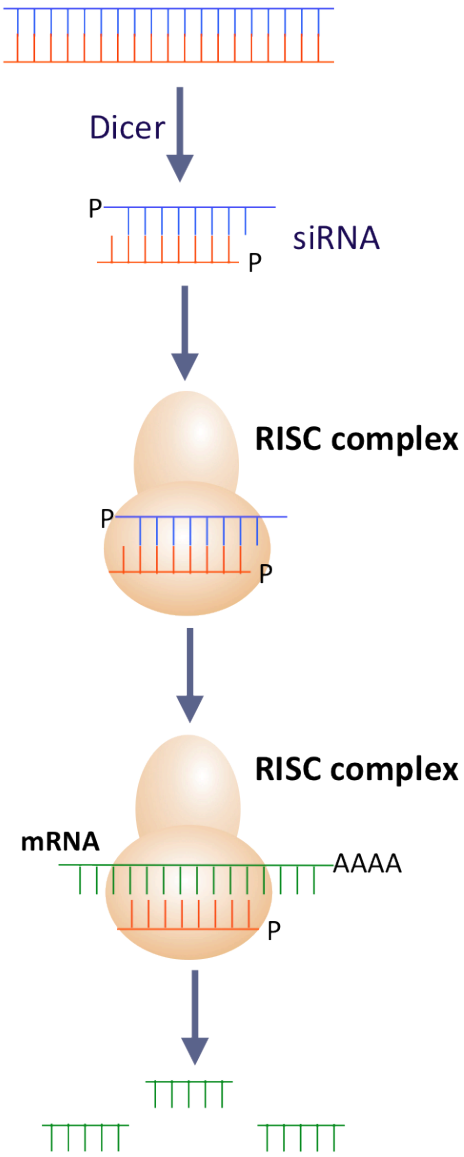

Figure 1.20. Mechanisms of generation of short RNAs. Long double-stranded exogenous RNA is processed into 21-23 nucleotides siRNA. The double-stranded siRNA is loaded onto the RISC (RNA-induced silencing complex) complex where the siRNA is unwind, leaving a single stranded guiding siRNA which guides the complex by sequence complementarity to target mRNA and the RISC complex then cleaves the target mRNA.

1.3.1 Application of siRNA technology in mammalian cells

Initial efforts to apply the siRNA technology in mammalian cell lines in culture were precluded by non-specific activation of antiviral responses triggered by dsRNA longer than 30 nucleotides (Yang *et al.*, 2001; Elbashir *et al.*, 2001; Caplen *et al.*, 2001). However, the Tuschl group has managed to circumvent this issue by using a chemically synthesised 21-nucleotide long siRNA duplex (instead of a long dsRNA), which is able to induce gene silencing upon introduction of the siRNA duplex into mammalian cells (Elbashir *et al.*, 2001). This represents an important milestone in functional annotation of genes in the mammalian systems, in which functions of many genes remain to be elucidated and has initiated efforts to perform large-scale screening experiments to reverse genetic analysis (Ito *et al.*, 2005). The Bernards group and Elledge group were among the first to perform large-scale mammalian siRNA screening experiments (Paddison *et al.*, 2004; Berns *et al.* 2004). The Bernards group developed a short-hairpin RNA (shRNA) library against 7914 human genes and subsequently performed a screen to identify genes involved in the p53-dependent proliferation arrest (Berns *et al.*, 2004). They created a primary human fibroblast cell line expressing murine telomerase catalytic subunit (TERT) and temperature sensitive allele of SV40 large T antigen, and a temperature shift to 39˚C is able to induce proliferation arrest in a p53-dependent manner. Using this cell system, they performed a RNAi screen and identified six genes, which when silenced, could suppress the proliferation arrest induced by temperature shift. One of those genes is p53 itself, underscoring the quality of the shRNA library generated.

As opposed to genome wide siRNA libraries which are laborious and cost-intensive to generate, there are also rationally designed siRNA libraries comprising of siRNAs targeting a particular protein family (Cullen and Arndt, 2005). For example, Aza-Blanc and colleagues generated a synthetic siRNA library against 510 kinases (Aza-Blanc *et al.*, 2003). Using the library, they have identified several genes which are involved in TNF-related apoptosis inducing ligand (TRAIL)-induced apoptosis. Our laboratory also adopts the similar strategy. We utilise a siRNA library against the human DUBs to reveal functional roles of this class of enzymes.

1.3.2 Off-target effects of siRNAs

During its early applications, siRNAs were thought to be highly specific, that target genes, were shown to be specifically silenced by complementary siRNAs only (Caplen et al, 2001; Tuschl *et al.*, 1999). Moreover, a single mismatch at the cleavage site of the siRNA is sufficient to abolish silencing effect (Elbashir, 2001; Amarzguioui *et al.*, 2003; Holen et al, 2002). However, with burgeoning experimental data, there was a growing notion of alternative mechanisms of transcript targeting by siRNAs (Jackson and Linsley, 2010). Several genome-scale expression profiling studies following siRNA depletion of specific target genes revealed the off-target activity of siRNAs (Chi *et al.*, 2003; Jackson *et al.*, 2003; Semizarov *et al.*, 2003). Following silencing the same gene using different siRNAs, it was revealed by microarray profiling that each siRNA exhibited a unique, sequence-dependent signature (Jackson *et al.*, 2003). Sequence analysis of the transcripts, which were non-specifically silenced revealed that their 3' UTRs were complementary to the 5' end of the siRNA guide strand. Notably, 8 complementary nucleotides to the siRNA are sufficient to cause an off-target effect. Subsequent works then established that sequence complementarity to the 5' end of the siRNA guide strand is responsible for off-target gene silencing events (Jackson *et al.*, 2006; Birmingham *et al.*, 2006). It was later revealed that the phenomenon of off-target silencing shares the same regulatory mechanism with microRNA-mediated gene silencing (Jackson and Linsley, 2010). As is the case for siRNA, the hundreds of mRNA targets regulated by a microRNA has, in their 3' UTR, a sequence, which is complementary to the 5' end of the microRNA guide strand, known as the "seed region" (Wightman *et al.*, 1999; Lai, 2002). The seed region plays pivotal role in target recognition by the microRNAs and is highly conserved among the metazoan microRNAs (Lewis *et al.*, 2003; Lim *et al.*, 2003). In other words, the 5' end of the guide strand of siRNAs is acting as a seed region motif to target and inhibit a wide array of genes.

To distinguish between on- and off- targets, a common approach is to silence the same gene with multiple siRNAs (Jackson and Linsley, 2010). Several siRNAs yielding the same phenotypic output is a strong indication of an on-target effect.

Another approach to confirm a phenotype resulting from the siRNA depletion of a gene is to rescue the phenotype by reintroducing into the cell a siRNA-resistant version of the same target gene (Cullen, 2006; Echeverri *et al.*, 2006). However, these experiments are technically challenging and interpretation of the experimental results can be difficult due to issues of transfection efficiency and expression levels (Jackson and Linsley, 2010). Endonuclease-prepared esiRNA (esiRNA) represents another solution to the problem of off-target effects (Kittler and Buccholz, 2005) (discussed in more detailed in Chapter 5). Essentially, esiRNA comprises of a complex mixture of siRNA generated by random cleavage of a long dsRNA and it has been demonstrated that an increasing complexity of pool of siRNAs minimises off-target effects (Kittler *et al.*, 2004).

1.4 Aims of this study

As discussed in the previous two sections of this chapter, both the core components of AJ, E-cadherin and β-catenin, are implicated in cancer development, where E-cadherin is acting as a tumour suppressor and dysregulation of β-catenin stability is oncogenic. Along with this notion, there is emerging evidence demonstrating ubiquitylation as one of the key post-translational modifications involved in regulation of these two proteins, and that the aberration in the ubiquitin-related machineries can also culminate in cancer development. Among the multiple aspects of ubiquitin-mediated cellular processes, our laboratory has a major interest in understanding the roles of the human deubiquitylases, the functions of most of which are yet to be annotated. One common approach we undertake to annotate function of the deubiquitylases is siRNA DUB library screen.

The aims of my research project are:

- to develop an assay to robustly assess E-cadherin status in a human cancer cell line.
- to utilise this assay as a readout for a siRNA DUB library screen to identify DUBs involved in E-cadherin regulation.

- to identify DUBs involved in regulation of E-cadherin at steady state conditions following siRNA DUB library screen.
- to identify DUBs involved in regulation of β-catenin at steady state conditions following siRNA DUB library screen.
- to decipher the mechanisms by which E-cadherin or β-catenin is regulated by the DUBs identified.
- to characterise the DUBs involved in E-cadherin or β-catenin regulation.
- to develop an endonuclease-prepared siRNA (esiRNA) DUB library as an alternative tool to siRNA.
- to perform quality control and characterise the esiRNA DUB library generated.

Chapter 2

Materials and Methods

2.1 Molecular Biology

2.1.1 Materials

Deoxynucleotide mix (100mM, ♯200415) and PfuUltra Hotstart DNA Polymerase (♯600390) were obtained from Stratagene (La Jolla, CA, USA). Taq DNA Polymerase (BIO-21040) was obtained from Bioline (London, UK). Gateway BP and LR clonase® II enzyme mix (♯11789-020 and ♯11791-100 respectively), DH5α subcloning efficiency cells (♯18265-017), S.O.C. medium (♯15544-034) and electrophoresis grade agarose were from Invitrogen (Paisley, UK). XL1-Blue supercompetent cells (♯200236) were obtained from Agilent Technologies (Santa Clara, CA, USA). TAE buffer was obtained from National Diagnostics (Hull, UK). Luria-Bertani broth (♯LAB191) was obtained from Lab M (Bury, UK). Miniprep (♯27106), HiSpeed Maxiprep (♯12633), RNeasy Mini (♯74106), gel extraction (♯28604) and PCR purification (♯28704) kits were all obtained from QIAGEN (Crawley, UK). Quick Ligation Kit (♯M2200S), 1kDA (♯N3232), 100bp (♯N3231), and Low Molecular Weight DNA ladder (♯N3233) were obtained from New England Biolab (Herts, UK). PCR Nucleotide Mix (♯C1441), RNasin® plus RNAse inhibitor (♯N2611) and nuclease free water (♯P1193) were obtained from Promega (Madison, WI, USA). M-MulV Reverse transcriptase (♯EP0352) was obtained from Thermo Scientific Inc (Waltham, MA, USA). All primers were ordered from Eurofins MWG Operon (Ebersberg, Germany). All other chemicals were obtained from Sigma Aldrich (Poole, UK) unless otherwise stated.

2.1.2 Polymerase Chain Reaction (PCR)

PCR (Mullis et al., 1986) were performed to amplify DNA sequences for subsequent cloning, to detect presence of plasmid containing genes of interest transformed into bacteria, and to prepare template for *in vitro* transcription during esiRNA production. In this study, two DNA polymerases were used, namely Taq DNA Polymerase (Bioline) and PfuUltra Hotstart DNA Polymerase (Stratagene). Taq DNA polymerase was used for bacteria colony PCR (B-PCR) to identify colonies positive for desired plasmids. PfuUltra Hotstart DNA polymerase was used in PCR reactions to amplify full length or fragments of desired genes. In the latter PCR reactions, plasmids were used as templates unless otherwise specified. PCR reaction mixtures were set up as summarised in Table 2.1 and Table 2.2, in 0.5ml Thermowell^{TM} tubes (Corning) at room temperature. PCR reactions were then allowed to run according to the thermal cycle programmes as shown in Table 2.3 and Table 2.4.

Table 2.1. PCR reaction mixtures for Taq DNA Polymerase.

Reagents	**B-PCR**
NH_4 Buffer (10X)	2µl
$MgCl_2$ (50mM)	0.8µl
dNTPs (100mM)	0.4µl
Fwd Primer (10µM)	0.5µl
Rev Primer (10µM)	0.5µl
Taq Polymerase (5U/µl)	0.1µl
NaOH (20mM)	-
Water	15.7µl
Total Volume	20µl

Table 2.2. PCR reaction mixtures for PfuUltra Hotstart DNA Polymerase.

Reagents	
PfuUltra HF Reaction Buffer (10X)	5µl
Template (25ng/µl)	1µl
dNTPs (100mM)	0.5µl
Fwd Primer (10µM)	1.25µl
Rev Primer (10µM)	1.25µl
PfuUltra HS DNA Polymerase (2.5U/µl)	1µl
Water	40µl
Total Volume	50µl

Table 2.3. Thermal cycler programme for Taq DNA polymerase.

Segment	Step	No. of cycles	Temperature	Duration
1	Initial Denaturation	1	95˚C	5 min
2	Denaturation	39	95˚C	1 min
	Annealing		T_m-5˚C	1 min
	Extension		72˚C	1 min /kb
3	Final Extension	1	72˚C	5 min

Table 2.4. Thermal cycler programme for PfuUltra Hotstart DNA polymerase.

Segment	Step	No. of cycles	Temperature	Duration
1	Initial Denaturation	1	95˚C	5 min
2	Denaturation	34	95˚C	1 min
	Annealing		T_m-5˚C	1 min
	Extension		68˚C	1 min /kb
3	Final Extension	1	68˚C	5 min

The B-PCR reactions were performed using primers specific to the vector pCR4TOPO or pDONR223. The primer sequences are shown in Table 2.5.

Table 2.5. Primers used for bacterial colony PCR. Sequences are in 5' to 3' direction.

Target	Primer Name	Primer Sequence
pCR4TOPO	M13F	CTGGCCGTCTTTTAC
pCR4TOPO	M13R	CAGGAAACAGCTATGAC
pDONR223	GateF2	TCGCGTTAACGCTAGCATGGAT
pDONR223	GateR1	GTAACATCAGAGATTTTGAGACAC

The primers used to amplify full length or fragments of USP38 were flanked by GATEWAY™ recombination sequence on their 5' ends. This allows shuttling of the amplified regions into a Gateway™ entry vector and subsequently into any of the Gateway™ expression vectors. An additional restriction site for specific endonuclease is also incorporated at the 3' end of the Gateway™ recombination sequence. This enables transfer of the amplified regions into expression vectors, which are not Gateway™ compatible, by simple DNA ligation reaction. At the 3' end of the restriction site is the Kozak sequence, ACC, which favours initiation of translation at the ATG start codon in higher eukaryotes (Kozak, 1987; Kozak, 1990). Table 2.6 shows the "Gateway" primers used in this study.

Table 2.6. List of "Gateway" primers used to amplify full length or regions of USP38. GATEWAY™ compatible attB flanking sequences are highlighted in bold and added endonuclease restriction sites are highlighted in red. Kozak sequence inserted into the forward primers is shown underlined and stop codon in reverse primers are highlighted in blue.

Primer Name	Primer Sequence
JLW_USP38_F_GW_EcoRI	**GAATTCACAAGTTTGTACAAAAAAGCAGGCTG** **G**GAATTCACCATGGACAAGATCCTGGAGGGC C
JLW_USP38_F_GW_HindIII	**GAATTCACAAGTTTGTACAAAAAAGCAGGCTG** **G**AAGCTTACCATGGACAAGATCCTGGAGGGCC
JLW_USP38_R_GW_SalI	**GGGGACCACTTTGTACAAGAAAGCTGGGTG**GT CGACTTAAAATACGAGTCTGCCAACTGTAT
JLW_USP38_F_1321_GW_BamHI	**GAATTCACAAGTTTGTACAAAAAAGCAGGCTG** **G**GGATCCACCGAAACTGGGAAAACTGGTC
JLW_USP38_R_2838_GW_SalI	**GGGGACCACTTTGTACAAGAAAGCTGGGTG**GT CGACTTATACAAAAGCACATAAGCTGTG
JLW_USP38_R_1323_GW_SalI	**GGGGACCACTTTGTACAAGAAAGCTGGGTG**GT CGACTTATTCAGATTTTCCAGAAAGTCTAG
JLW_USP38_F_1321_GW_EcoRI	**GAATTCACAAGTTTGTACAAAAAAGCAGGCTG** **G**GAATTCACCGAAACTGGGAAAACTGGTC
JLW_USP38_R_1314_GW_SalI	**GGGGACCACTTTGTACAAGAAAGCTGGGTG**GT CGACTTATCCAGAAAGTCTAGACAAGCAAG

2.1.3 DNA agarose gel electrophoresis

Agarose gels (between 0.8 – 2%) were prepared by adding electrophoresis grade agarose powder to 1X TAE buffer (40mM Tris Acetate, 1mM Na_2EDTA). The mixture was then heated using a microwave to fully dissolve the agarose. The agarose solution was allowed to cool down at room temperature for about 15 minutes, after which ethidium bromide was added to a final concentration of 0.5µg/ml. The gel was then poured and allowed to set at room temperature. 10X DNA Sample buffer (5% w/v glycerol, 0.1mM EDTA, 0.04% bromophenol blue) was added to DNA sample at a ratio of 1:10 (i.e. 1µl 10X DNA sample buffer to 10µl DNA sample) and mixed well by gentle vortexing. The samples were then loaded onto the gel, along with 10-15µl per lane of DNA ladder standard, and resolved in TAE buffer, in a Fisherbrand horizontal midi electrophoresis tank (Fisher Scientific, Loughborough, UK), at 110V to 120V for an hour. DNA bands were visualized using an ultraviolet light source.

2.1.4 Bacterial transformation

For each transformation reaction, 50µl of competent cells was thawed on ice and transferred to a polypropylene tube. Less than 100ng of DNA was added to the cell suspension and incubated on ice for 20 minutes. The cell suspension was then heat-shocked at 42˚C in a water bath for 45 seconds and placed immediately on ice after that for 2 minutes. 200µl of S.O.C. (super optimal broth with catabolite repression) medium was then added to the cell suspension and the bacteria were shaken at 250rpm for 1 hour. Bacterial cells were then spread on selective LB agar plates and incubated at 37˚C overnight.

2.1.5 Site-directed mutagenesis

Site-directed mutagenesis was performed to generate the catalytically inactive mutant of USP38 and BAP1, and siRNA resistant plasmid of BAP1. The reaction was carried out using the QuickChange site-directed mutagenesis kit (Agilent Technologies) and PfuUltra Hotstart DNA Polymerase. The mutagenesis reaction was performed using pairs of complementary primers covering the region to be mutagenised (single or multiple base(s) change) as shown in Table 2.7. The template used was either pDONR223-USP38 or pDONR223-BAP1. Reaction mixtures were set up in 0.65ml Thermowell^{TM} at room temperature as shown in Table 2.8 and reactions were allowed to run according to the thermal cycle programmes as shown in Table 2.9. At the end of the PCR reactions, 1µl of DpnI restriction enzyme was added and the mixture incubated at 37˚C for 1 hour. DpnI restriction enzyme recognises the DNA sequence GATC with methylation of the Adenine nucleotide. The purpose of this step is to remove the wildtype plasmids purified from bacteria, which were modified by methylation. The newly generated plasmids by PCR reactions were not methylated and therefore were not recognised by the DpnI enzyme. After that, 2-5µl of the final reacted mixtures were used to transform XL1Blue cells. Plasmids were then purified from positive transformants using MiniPrep Kit (Qiagen) and tested by restriction digest. Candidate mutated plasmids were then sent for sequencing and glycerol stock was made.

Table 2.7. Site-directed mutagenesis primers. Mutated nucleotide residues are shown in upper case letters.

Target	Primer Name	Primer Sequence	Base change	Amino acid change
USP38	JLW_USP38_C454A_F	cctaggaaatacaGCttatatg aacagtg	TGt → GCt	C454A
	JLW_USP38_C454A_R	cactgttcatataaGCtgtattt cctagg		
BAP1	BAP1_C91S_Fwd	ctgatacccaactcttCtgcaa ctcatgccttgctg	tGt → tCt	C91S
	BAP1_C91S_Rev	cagcaaggcatgagttgcaGa agagttgggtatcag		
BAP1	BAP1_A95D_Fwd	gtgcaactcatgActtgctgag cgtgctcc	gCc → gAc	A95D
	BAP1_A95D_Rev	ggagcacgctcagcaagTcat gagttgcac		
BAP1	BAP1_siRes_OL3_2_F	gtccccgctggtgTtAgaGgc Gaacagggcccctgcag	CtGgaAg cA → TtAgaGg cG	-
	BAP1_siRes_OL3_2_R	ctgcaggggccctgttCgcCtc TaAcaccagcggggac		

Table 2.8. Reaction mixtures for site-directed mutagenesis.

Reagents	
PfuUltra HF Reaction Buffer (10X)	5µl
Template (25ng/µl)	1µl
dNTPs (100mM)	0.5µl
Fwd Primer (10µM)	1.25µl
Rev Primer (10µM)	1.25µl
PfuUltra HS DNA Polymerase (2.5U/µl)	1µl
Water	40µl
Total Volume	50µl

Table 2.9. Thermal cycler programme for site-directed mutagenesis.

Segment	No. of cycles	Temperature	Duration
1	1	95˚C	30 sec
2	15	95˚C	30 sec
		55˚C	1 min
		68˚C	2 min /kb

2.1.6 TOPO cloning

TOPO cloning reaction allows direct insertion of PCR products into the pCR4®Blunt-TOPO vector, catalysed by blunt end topoisomerase. ORFs of interest were first amplified by PCR and resolved on a DNA agarose gel. The PCR product was then extracted from the gel using the gel extraction kit from QIAGEN and used for TOPO cloning reaction. TOPO cloning reactions were set up on ice as shown in Table 2.10 and were then incubated at room temperature for 15 minutes. After that, 6µl of reaction mixture was used to transform MACH1 competent cells. Cells were spread on LB agar plates containing Ampicillin (100mg/ml) or Kanamycin (10mg/ml) and were incubated overnight at 37˚C. On the next day, bacterial colony PCR (see section 2.1.2) was performed, using M13F and M13R primers, for a minimum of 4 bacterial colonies to screen for positive transformants. Mini-prep cultures were then set up for positive transformants and DNA was purified from bacterial culture by QIAGEN MiniPrep Kit. The purified DNA was subsequently subjected to diagnostic test digest and DNA, which was tested positive was then sent for sequencing. Glycerol stocks of bacterial colonies transformed with sequence-verified pCR4®Blunt-TOPO constructs were made as described in section 2.1.9.

Table 2.10. Reaction mixture for TOPO cloning reaction.

Reagent	Volume
pCR4®Blunt-TOPO vector	1 µl
Salt solution	1 µl
PCR product	4 µl
Total volume	6 µl

2.1.7 Gateway cloning

Gateway cloning (Hartley et al., 2000) was performed to allow shuffling of open reading frames into plasmids of interest.

2.1.7.1 Gateway BP cloning reaction

PCR products of ORFs flanked by *attB* recombination sites were inserted into the pDONR223 entry vector by BP recombination reaction. BP reactions were carried out using a specific insert to entry vector ratio, which was calculated using the following formula:

$$\text{Amount of insert (ng)} = \frac{125 \times \text{size of ORF} \times 660}{10^6}$$

BP reaction tubes were set up on ice as shown in Table 2.11, and incubated at 25˚C for 1.5 hour. After this, 1µl of Proteinase K (from BP reaction kit, Invitrogen) was added to the reaction mixture and incubated for 10 minutes at 37˚C to digest the BP clonase enzyme. 6µl of reaction mixture was then used to transform 50µl of XL1-Blue competent cells. These cells were spread on LB agar plates containing Spectinomycin (100µg/ml) and the plates were incubated overnight at 37˚C. On the next day, bacterial colony PCR (see section 2.1.2) was performed, using GateF2 and GateR1 primers, for a minimum of 5 bacterial colonies to screen for positive transformants. Miniprep cultures were then set up for positive transformants and DNA was purified from bacterial culture by QIAGEN MiniPrep Kit. The purified DNA was subsequently subjected to diagnostic test digest and DNA, which was tested positive was then sent for sequencing.

Table 2.11. **Reaction mixture for BP cloning reaction.**

Reagent	**Volume**
BP clonase II enzyme mix	2 µl
pDONR223 vector (300ng/µl)	1 µl
PCR product or linearized pCR4Blunt-TOPO construct	x µl
Sterile water	y µl
Total volume	10 µl

2.1.7.2 Gateway LR Cloning Reaction

Gateway LR reaction allows the shuttling of DNA insert from entry vector (pDONR223) to destination vectors by recombination reaction, catalysed by LR clonase enzyme. LR cloning reactions were set up on ice as shown in Table 2.12 and the reaction tubes were incubated at 25˚C for 1.5 hour. Following the incubation period, 1µl of Proteinase K (Invitrogen) was added to the reaction mixture and incubated at 37˚C for 10 minutes.

Table 2.12. **Reaction mixture for LR cloning reaction.**

Reagent	**Volume**
BP clonase II enzyme mix	2 µl
pDONR223 vector (300ng/µl)	1 µl
PCR product or linearized pCR4Blunt-TOPO construct	x µl
Sterile water	y µl
Total volume	10 µl

6µl of reaction mixture was then used to transform 50µl of DH5α competent cells. These cells were spread on LB agar plates containing Kanamycin (10mg/ml) and the plates were incubated overnight at 37˚C. On the next day, at least 3 colonies were picked to set up miniprep cultures and DNA was purified from the bacterial culture on the following day by QIAGEN MiniPrep Kit. The purified DNA was subsequently subjected to diagnostic test digest and colonies carrying plasmids positive for test digest were used to make glycerol stocks.

2.1.8 Restriction Digest

Restriction digest reactions were set up as shown in Table 2.13 and incubated at the optimal temperature (mostly 37˚C) for the restriction endonucleases used and for 1 hour. Typically, 800ng of DNA sample was used for diagnostic test digest and 5µg of DNA sample was used for restriction cloning.

Table 2.13. Reaction mixture for restriction digest.

DNA sample	x µl
Restriction endonuclease	1.0 µl
Reaction buffer	1.5 µl
BSA	1.5 µl
Sterile water	y µl
Total volume	15 µl

2.1.9 Restriction Cloning

Restriction cloning was performed, using the Quick Ligation kit from NEB. Insert was excised from an original vector using a pair of restriction endonucleases (see section 2.1.7), and the destination vector to which the insert is going to be ligated was also digested with a compatible pair of restriction endonucleases (i.e. the sticky ends generated following restriction digest are complementary to each other). Typically, 100ng of restriction-digested insert was incubated with vector at a molar ratio of 1:3. Ligation reactions were set up as shown in Table 2.14 on ice, and were incubated at room temperature for 10 minutes. 6µl of reaction mixture was then used to transform 50µl of DH5α competent cells. These cells were spread on LB agar plates containing Kanamycin and the plates were incubated overnight at 37˚C. On the next day, at least 4 colonies were picked to set up miniprep cultures and DNA was purified from the bacterial culture on the following day by QIAGEN MiniPrep Kit. The purified DNA was subsequently subjected to diagnostic test digest and colonies carrying plasmids positive for test digest were used to make glycerol stocks.

Table 2.14. Reaction mixture for ligation reaction.

Restriction digested insert (100ng)	x µl
Restriction digested vector	y µl
2X Quick Ligase Buffer	5 µl
Quick Ligase	0.5 µl
Sterile water	z µl
Total volume	10 µl

2.1.10 Glycerol Stock

To make glycerol stocks of transformed bacteria, 5ml of overnight bacterial culture in LB medium (inoculated from a single bacterial colony) was spun down at 4000rpm in a centrifuge. Supernatant was then discarded and the bacterial pellet was resuspended in 40% glycerol in LB. The glycerol stock was then stored at -80˚C.

2.1.11 Reverse Transcription

mRNA was harvested from a monolayer of adherent cells using QIAGEN RNeasy Kit as per manufacturer guidelines. The concentration of the RNA extract was measured using a NanoDrop Spectrophotometer ND1000 at a wavelength of 260nm. 1µg of mRNA was then diluted with sterile water to a volume of 10µl, followed by addition of 1µl of oligo DT primers. The mixture was then incubated at 70˚C for 5 minutes to allow priming of the oligo DT primers to poly-A tails of mRNA. In the mean time, a reaction buffer mixture was prepared as shown in Table 2.15. 8µl of this reaction buffer was then added to the tube with diluted mRNA and oligo DT primers, and the tube was incubated at 37˚C for 5 minutes. After that, 1µl of M-MuLV reverse transcriptase was added to the tube and the reaction mixture was incubated at 42˚C for 1 hour, followed by a 10-minute incubation at 70˚C. The resulting cDNA was diluted to 100µl with sterile water.

Table 2.15. Reaction buffer mixture for reverse transcription.

5X reverse transcription buffer	4 µl
PCR nucleotide mix	2 µl
RNasin	0.5 µl
Nuclease free deionised water	1.5 µl
Total volume	8 µl

2.1.12 Quantitative Real-Time PCR (QPCR)

QPCR (Higuchi et al., 1992) were performed to determine the level of CDH1 mRNA and CTNNB1 mRNA following knockdown of USP38 and BAP1 respectively. QPCR reaction mixture was set up using DyNAmo™ HS SYBR® Green qPCR kit as per manufacturer's instruction. QPCR reactions were allowed to run on iQ5 real-time PCR detection system (Bio-Rad) according to the following protocol:

Table 2.16. QPCR setting.

Segment	Step	No. of cycles	Temperature	Duration
1	Initial Denaturation	1	95˚C	15 min
2	Denaturation	40	95˚C	10 sec
	Annealing		60˚C	30 sec
	Extension		72˚C	30 sec
3	Melting curve analysis	81 repeats	55˚C to 95˚C	30 sec for every 0.5˚C

Fluorescence measurement of each reaction well was taken at real time at the end of each cycle of the second segment. The quantity of the target gene transcript was normalised against that of the reference gene, which was actin.

2.2 Cell Biology

2.2.1 Materials

Oligofectamine™ transfection reagent (♯12252-011), Lipofectamine™ LTX reagent (♯15338-100) and all cell culture reagents were obtained from Invitrogen unless otherwise stated (Paisley, UK). HiperFect transfection reagent (♯301705) was obtained from QIAGEN (Crawley, UK). GeneJuice® transfection reagent (♯70967-3), Folimycin (♯344085) and Lactacystin (♯426100) were obtained from EMD Millipore (Darmstadt, Germany). All other chemicals were obtained from Sigma Aldrich (Poole, UK) unless otherwise stated. All plasticware was obtained from Corning Inc (NY, USA). HGF was received as a kind gift from George Vande Woude (Van Andel Research Institute, Grand Rapids, MI).

2.2.2 Cell Culture

Unless otherwise stated, all human cell lines (MCF7, SW480, HEK293T, A549 and DU145) were maintained in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, Paisley, UK), supplemented with 10% heat-inactivated foetal bovine serum (FBS), 0.1mM MEM non-essential amino acids (Gibco), 100 units ml^{-1} Penicillin (Gibco) and 100 units ml^{-1} Streptomycin (Gibco). All cells were incubated at 37˚C, 5% CO_2. For maintenance, confluent layer of cells (90-95%) were detached from dish using 0.05% Trypsin with EDTA (Gibco) and were split at the following densities for every two days:

Table 2.17. Splitting density of different cell lines.

Cell Line	Split Density
MCF7	1 in 3
A549	1 in 5
SW480	1 in 3
HEK293T	1 in 8
HeLa	1 in 5
DU145	1 in 5

siRNA transfection of A549, MCF7, HeLa and HEK293T cells was performed using oligofectamine. Typically, cells were seeded into wells of 6-well plate and incubated overnight at 37˚C. On the next day, the cells were washed with PBS and 800µl of no addition DMEM (DMEM without FBS) was added. 2µl of 20µM siRNA oligos was first added to 180µl of OptiMEM (Invitrogen). In a separate tube, 3µl of oligofectamine transfection reagent was added to 13µl of OptiMEM, and mixed by vortexing. Both solutions were left to stand at room temperature for 5 minutes. After that, the solutions were mixed together, vortexed and left stand at room temperature for 20 minutes. The 200µl of transfection mixture was then added dropwise to each well and the resulting concentration of siRNA oligos is 40nM. The cells were then incubated at 37˚C for 4-6 hours, after which the medium was replaced with full DMEM (DMEM containing 10% FBS). Cells were normally lysed or fixed 72 hours post-transfection.

siRNA transfection of SW480 cells was performed using HiperFect transfection reagent (QIAGEN). Typically, cells were seeded into wells of 6-well plate and incubated overnight at 37˚C. On the next day, the cells were washed with PBS and 800µl of no addition DMEM was added. 2µl of 20µM siRNA oligos was first added to 183µl of OptiMEM (Invitrogen), mixed well by vortexting and left stand at room temperature for 5 minutes. After that, 15µl of HiperFect transfection reagent was added and the transfection mixture was mixed by vortexing and left stand at room temperature for 20 minutes. The 200µl of transfection mixture was then added dropwise to each well and the resulting concentration of siRNA oligos is 40nM. The cells were then incubated at 37˚C for 4-6 hours, after which the medium was replaced with full fresh DMEM. Cells were normally lysed or fixed 72 hours post-transfection.

2.2.4 DNA Transfection

GeneJuice was used for transient DNA transfection of HeLa and HEK293T and LipoFectamine LTX was used for transient DNA transfection of MCF7. Typically, cells were seeded into wells of 6-well plate and incubated overnight at 37˚C. On the

following day, the medium in the well was replaced by 1ml fresh full DMEM. To prepare the transfection mixture, typically 1µg of plasmid was added to 150µl (for GeneJuice transfection) or 500µl (for Lipofectamine LTX transfection) of OptiMEM and mixed well by gentle vortexing. Then, 3µl of GeneJuice or 6µl of Lipofectamine LTX transfection reagent was added, followed by vortexing, and standing at room temperature for 20 minutes. The transfection mixture was subsequently added dropwise into the corresponding well. Cells were then incubated at 37˚C and lysed or fixed 24 hours post-transfection.

2.2.5 Cell Aggregation Assay

A cell aggregation assay (Nola et al., 2012) was performed to assess ability of MCF7 cells to aggregate following knockdown of USP38 (see chapter 3, section 3.2.5). Typically, monolayer of adherent cells were washed twice with warm PBS, followed by addition of 2ml (for a well of a 6-well plate) of 2mM EDTA in PBS. Cells were then incubated at 37˚C for about 15 minutes until they detached from the plate. 4 ml of full DMEM was added and cells were pipetted up and down 10 times to separate cells from each other. Then, cell number was counted using a haematocytometer and cells were resuspended to a density of 5 x 10^5 cells per ml. The lid of a 6cm dish was inverted and several 20µl droplets of the cell suspension were then pipetted onto the inner side of the lid, and the dish was filled with 3ml of sterile water. The lid was quickly put on top of the dish, so that the cells would fall to the tip of the droplet by gravity, and incubated at 37˚C for 2 hours. A phase contrast image of each droplet was taken at 1 hour intervals.

2.3 Protein Biochemistry

2.3.1 Materials

BL21 expression competent cells (♯200131) were obtained from Agilent Technologies (Santa Clara, CA, USA). Protein G Sepharose (♯P3296) and Protein A Sepharose (♯P9424) were obtained from Sigma Aldrich. BCA protein assay kit

(♯23225) and Glutathione-agarose beads (♯15160) were obtained from Pierce Biotechnology (Rockford, IL, USA). Bio-Rad Protein Assay Kit (♯500-0002) was obtained from Bio-Rad (Hemel Hempstead, UK). Protran Nitrocellulose Membrane (♯B3-0042), ProtoGel Resolving Buffer (♯B9-0010), ProtoGel Stacking Buffer (♯B9-0014), SDS PAGE Tank Buffer, 10X (♯B9-0032), ProtoGel (♯A2-0072) were obtained from GeneFlow (Elmhurst, UK). NuPAGE® 4-12% Bis-Tris gels (♯NP0321BOX, ♯NP0323BOX, ♯NP0326BOX, ♯WG1402BOX) and NuPAGE® MOPS 20X SDS Running Buffer (♯NP-0001-02) were obtained from Invitrogen (Paisley, UK). All other chemicals were obtained from Sigma-Aldrich (Poole, UK) unless otherwise stated.

2.3.2 Bacterial Protein Purification

GST-tagged RNAse III protein was purified from BL21 expression competent cells. BL21 cells were first transformed with pGEX2T-RNAse III (Section 2.1.4) and a colony tested positive for the plasmid was picked and grown overnight in 50ml of LB medium (+ Ampicillin) at 37˚C and with shaking at 250rpm. The overnight culture was used to set up a 2 x 500ml LB (+ Ampicillin) at a 1 in 20 dilution, and was allowed to grow until the optical density (600nm) reached 0.8. Protein expression of GST-RNAse III was then induced with the addition of IPTG at a final concentration of 0.4mM for 4 hours at 37˚C.

Following protein expression induction, bacteria was pelleted at 4000rpm and 4˚C for 15 minutes. The bacterial pellet was resuspended in 50ml of lysis buffer (150mM NaCl in PBS plus bacterial protease inhibitor cocktail), transferred to a 50ml Falcon tube and centrifuged at 4000rpm. The supernatant was discarded and the pellet resuspended in 25ml of lysis buffer. Lysozyme was then added to the cell suspension, at a concentration of 1mg/ml, and the cell suspension was incubated on ice for 15 minutes.

Bacterial cells were then lysed using a sonicator probe, 4 times 30 seconds at a power of 10 and on ice. The resulting crude lysate was spun, in Optima

ultracentrifuge, at 55,000g for 30 minutes. The supernatant was collected and the pellet was resuspended to the original volume. The supernatant was then combined with 1.3ml of 75% (v/v) slurry glutathione-agarose beads (prewashed three times using PBS) in a 50ml Falcon tube, and was incubated at 4˚C on a rotating wheel for 2 hours. The agarose beads were then spun down at 4000rpm for 5 minutes at 4˚C and the supernatant was retained. To wash the beads, they were resuspended in 20ml PBS (plus dithiothreitol) and spun down at 1000rpm, 4˚c for 1 minute. This step was repeated 3 times. Then, the protein was eluated 4 times using 1ml 50mM Tris-HCl (pH 8.0) plus 10mM reduced glutathione at room temperature for 10 minutes. Eluates were collected, combined to a final volume of about 4ml and dialysed against 2 x 1000ml of dialysis buffer (20mM HEPES-KOH, 100mM KCl and 0.2mM DTT; pH 7.6).

2.3.3 Cell Lysis for Protein Harvest

Monolayer of adherent cells were washed twice with ice-cold PBS on ice and lysed with NP40 lysis buffer (0.5% (w/v) NP40, 25mM Tris pH 7.5, 100mM NaCl, 50mM NaF) plus phosphatase inhibitor cocktail II (Sigma Aldrich) and mammalian protease inhibitor (Sigma Aldrich).

2.3.4 Protein Assay

The protein concentration of cell lysates was quantified using the BCA protein assay kit (♯23225, Pierce, UK) or the Bio-Rad Protein Assay Kit (♯500-0002) according to the manufacturer's instructions.

2.3.5 Co-immunoprecipitation

Cells were lysed as described in section 2.3.2. Protein concentration of lysate was determined (see section 2.3.4) and the lysates were adjusted to equal

concentration using lysis buffer in a fresh tube (typically, 500-800µg of protein for each co-immunoprecipitation). Antibody against protein of interest (1µg per µg of lysate) and 30µl of 50% (v/v) slurry protein-G (for primary antibodies raised in mice or sheeps) or protein-A (for primary antibodies raised in rabbits) Sepharose were added and incubated on a rotating wheel at 4˚C for 1-2 hour. Beads were pelleted at top speed in a bench top centrifuge and washed 3 times with YP-IP buffer (0.1% (w/v) NP40, 25mM Tris pH 7.5, 150mM NaCl), followed by a final wash in 10mM Tris pH 8.0. Proteins bound to beads were then retrieved by boiling the beads in 30µl 3X sample buffer for 5 minutes.

2.3.6 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Sodium dodecyl sulfate (SDS) Polyacrylamide gel electrophoresis was performed using the BioRad system or precast NuPAGE® 4-12% Bis-Tris gels (Invitrogen, Paisley, UK).

For the BioRad system, gels were poured at the following concentrations:

Table 2.18. Recipe for pouring two resolving gels.

Percentage Gel	8% (v/v)	10% (v/v)	12% (v/v)
Protogel (ml)	5.3	6.7	8
Protogel Resolving (ml)	5.2	5.2	5.2
Water (ml)	9.2	7.9	6.6
Temed (µl)	15	15	15
10% (w/v) APS (µl)	150	150	150
Total Volume (ml)	~20ml	~20ml	~20ml

Table 2.19. Recipe for pouring two 4% (v/v) stacking gel.

Protogel (ml)	1.3
Protogel Stacking (ml)	2.5
Water (ml)	6.1
Temed (µl)	50
10% (w/v) APS (µl)	10
Total Volume (ml)	~10ml

SDS-PAGE using precast NuPAGE® 4-12% Bis-Tris gel was performed at a constant voltage of 200V for 55 minutes (mini gel) or 65 minutes (midi gel). BioRad gels were typically run at 90V for 15 minutes followed by 135V for 70 minutes.

2.3.7 Western Blotting

Western blotting (Burnette, 1981) was performed to transfer proteins resolved on polyacrylamide gel to nitrocellulose membrane. Proteins were transferred from polyacrylamide gels onto 0.45µm Protran® nitrocellulose membrane (Geneflow) at a constant current of 0.8 – 0.9A for an hour. Following transfer, Ponceau-S stain (Sigma Aldrich) was used to stain the blot to visualise successful transfer of protein. The Ponceau-S stain was removed by washing with PBS. The blot was then blocked for a minimum of 1 hour using a suitable blocking buffer, which is compatible with the primary antibodies to be used. Typically, blots were blocked with 5% Marvel milk powder in TBS (20mM Tris, 137mM NaCl, pH 7.6) containing 0.1% Tween20 (hence TBS-T). Blots were then incubated with primary antibody for an indicated length of time (see Table 2.20). Blots were washed for 3 times 4 minutes using TBS-T, followed by incubation with fluorophore conjugated secondary antibody (see Table 2.21) at room temperature for an hour. Before acquisition of a blot image, the blot was washed for 3 times 4 minutes using TBS-T, followed by 1 time 4 minute wash using PBS.

Table 2.20. Primary antibodies for Western Blotting

Antibody name	Target protein	Immunogen	Manufacturer	Blocking buffer	Dilution
HECD1	E-cadherin	N-term	CRUK /Abcam (♯ab1416)	Milk	1:1000
EP700Y	E-cadherin	N-term	Abcam (♯ab40772)	Milk	1:2000
M168	E-cadherin	C-term	Abcam (♯ab76055)	Milk	1:1000
24E10	E-cadherin	Synthetic peptide around Pro780	Cell Signaling (♯3195)	Milk	1:1000
E247	β-catenin	Near N-term	Abcam (♯ab32572)	Milk	1:1000
8E7	β-catenin	HSGATTTAP	Millipore (♯05-665)	Milk	1:1000
Anti-β-catenin	β-catenin	aa 571 - 781	BD Transduction Lab (♯610154)	Milk	1:1000
Anti-α-tubulin	α-tubulin	C-term	Sigma (♯T6074)	Milk/BSA	1:10000
Anti-actin	Actin	SGPSIVHRKCF	Sigma (♯A2066)	Milk/BSA	1:2000
Anti-GFP	GFP	N/A	Ian Prior	Milk	1:1000
FK2	Ubiquitin	N/A	Millipore ((♯04263)	BSA	1:1000
USP8	USP8	aa 239 - 377	Sigma (♯HPA004869)	Milk	1:1000
USP9X	USP9X	N/A	Bethyl (♯A301-350)	Milk	1:1000
USP15	USP15	N/A	Bethyl (♯A300-923A)	Milk	1:1000
USP20	USP20	N/A	Bethyl (♯A301-189A)	Milk	1:1000
USP38	USP38	aa 700 - 750	Abcam (ab72244)	Milk	1:2000
BAP1	BAP1	aa 430 - 729	Santa Cruz (♯sc28383)	Milk	1:1000
A20	A20/TNFAIP3	Full length	Santa Cruz (♯sc52910)	BSA	1:1000

Table 2.21. Secondary antibodies for Western blotting. All antibodies were used in the same blocking buffer as primary antibodies.

Secondary Antibody	Manufacturer	Catalogue No.	Dilution
Donkey anti-mouse IRDye 800CW	LICOR Biosciences	926-32212	1:15000
Donkey anti-mouse IRDye 680CW	LICOR Biosciences	926-32222	1:15000
Donkey anti-rabbit IRDye 800CW	LICOR Biosciences	926-32213	1:15000
Donkey anti-rabbit IRDye 680CW	LICOR Biosciences	926-32223	1:15000
Donkey anti-sheep IRDye 800CW	LICOR Biosciences	926-32214	1:15000
Donkey anti-sheep IRDye 680CW	LICOR Biosciences	926-32224	1:15000

2.3.8 Immunofluorescence staining

Cells grown on coverslips were washed twice with room temperature PBS before incubation with 4% paraformaldehyde in PBS for 15 minutes. Cells were then washed twice using PBS and incubated with 50mM ammonium chloride in PBS for 15 minutes, followed by a further two rinses with PBS. Then, cells were permeabilised using 0.2% Triton (w/v)/PBS for 4 minutes and then blocked with 10% goat serum /PBS for 30 minutes. Cells were then incubated with primary antibody in 5% goat serum/PBS (see Table 2.22 for list of primary antibodies used) for 20 minutes, followed by 3 washes in PBS and incubation with secondary antibodies (see Table 2.23 for list of secondary antibodies used) in 5% goat serum/PBS for 20 minutes. Cells were then washed another 3 times in PBS, dipped into sterile deionised water and mounted onto slides with Mowiol plus 4',6-diamidino-2-phenylindole (DAPI) stain.

Table 2.22. Primary antibodies used for immunofluorescence staining. Blocking buffer and antibody incubation buffer is 10% and 5% goat serum / PBS respectively.

Antibody name	Target protein	Immunogen	Manufacturer	Dilution
HECD1	E-cadherin	N-term	CRUK /Abcam (♯ab1416)	1:1000
EP700Y	E-cadherin	N-term	Abcam (♯ab40772)	1:2000
M168	E-cadherin	C-term	Abcam (♯ab76055)	1:1000
24E10	E-cadherin	Synthetic peptide around Pro780	Cell Signaling (♯3195)	1:1000
E247	β-catenin	Near N-term	Abcam (♯ab32572)	1:1000
8E7	β-catenin	HSGATTTAP	Millipore (♯05-665)	1:1000
BAP1	BAP1	aa 430 - 729	Santa Cruz (♯sc28383)	1:1000
H4A3	LAMP-1	N/A	DSHB	1:50
CD63	CD63	N/A	Biodesign (♯N42768M)	1:100
EEA1 243/3	EEA1	N/A	Ian Mills	1:500

Table 2.23. Secondary antibodies used for immunofluorescence staining.

Secondary Antibody	Manufacturer	Catalogue No.	Dilution
Donkey anti-sheep AF488	Invitrogen	A11015	1:500
Donkey anti-sheep AF594	Invitrogen	A11016	1:500
Donkey anti-rabbit AF488	Invitrogen	A21206	1:500
Donkey anti-rabbit AF594	Invitrogen	A21207	1:500
Donkey anti-mouse AF488	Invitrogen	A21202	1:500
Donkey anti-mouse AF594	Invitrogen	A21203	1:500

2.4 esiRNA Production

An esiRNA DUB library was generated to facilitate further screening efforts in the laboratory.

2.4.1 Materials

HotStarTaq Master Mix Kit (♯203446) and QIAquick Gel Extraction Kit (♯28704) were obtained from Qiagen (Crawley, UK). MEGAscript T7 *in vitro* Transcription Kit was obtained from Ambion, Life Technologies (Paisley, UK). GST-RNASeIII was prepared by myself. Q-Sepharose Fast Flow (♯17-0510-01) was

obtained from GE Healthcare (UK). 1kDA (♯N3232), 100bp (♯N3231), and Low Molecular Weight DNA ladder (♯N3233) were obtained from New England Biolab (Herts, UK). Empty Bio-Spin Chromatography Column was obtained from Bio-Rad (Hemel Hempstead, UK). All other chemicals were obtained from Sigma Aldrich (Poole, UK) unless otherwise stated.

2.4.2 Polymerase Chain Reaction

PCR were performed to obtain PCR products as templates for the subsequent stage of *in vitro* transcription (IVT) reaction to produce double-stranded RNA (dsRNA). PCR products of esiRNA regions of the following DUBs were amplified from ORFs which were already cloned in the lab: CYLD, USP1, USP2A, USP3, USP4, USP5, USP6, USP7, USP8, USP9X, USP10, USP11, USP12, USP13, USP14, USP15, USP16, USP18, USP19, USP20, USP21, USP25, USP26, USP28, USP29, USP30, USP32, USP33, USP36, USP38, USP39, USP42, USP44, USP45, USP46, USP47, USP48, USP49, USP50, USP52, USP53, USP54, USPL1, BAP1, UCHL1, UCHL3, UCHL5, JOSD1, JOSD2, BRCC3, CSN5, CSN6, EIF3H, EIF3S5, MPND, MYSM1, PSMD14, PSMD7, STAMBP, STAMBPL1, OTUB1, OTUB2, OTUD3, OTUD4, OTUD5, OTUD6A, OTUD6B, OTUD7B, TNFAIP3, VCPIP1, YOD1, ZRANB1, GFP. The primer pairs used are flanked by T7 promoter sequence (for IVT) and are summarised in Appendix 2.1. These PCR reactions were set up using HotStarTaq Master Mix Kit (Qiagen) according to manufacturer's instruction and the reaction setting was as follows:

Table 2.24. Thermal cycler programme for PCR reactions to amplify esiRNA region.

Segment	Step	No. of cycles	Temperature	Duration
1	Initial Denaturation	1	95˚C	5 min
2	Denaturation	39	95˚C	1 min
	Annealing		T_m-5˚C	1 min
	Extension		72˚C	1 min /kb
3	Final Extension	1	72˚C	5 min

PCR products of esiRNA regions of the following DUBs were amplified from cDNA of MCF7 cells: DUB3, DUB4, USP17, USP22, USP24, USP27X, USP31, USP34, USP35, USP37, USP40, USP43, ATXN3, ATXN3L, PARP22, PRPF8, OTUD7A. cDNA was prepared from MCF7 RNA extract as described in section 2.1.11. The primer pairs used are flanked by T7 promoter sequence (for IVT) and are summarised in Appendix 2.1. The PCR reactions were set up using Pfu Ultra DNA polymerase as described in section 2.1.2 (Table 2.2). These PCR products were then cloned into pCR4®Blunt-TOPO vectors as described in section 2.1.6. The resulting plasmids were sequenced verified and a glycerol stock of each plasmid was made. The list of plasmids that were generated were summarised in Appendix 2.2.

2.4.3 *In vitro* transcription

PCR products that were obtained from ORFs of DUBs were analysed by DNA agarose gel electrophoresis and were purified using QIAquick gel extraction kit (Qiagen) according to manufacturer's instructions. DUBs for which their esiRNA region were inserted into pCR4®Blunt-TOPO vectors, the esiRNA region were released from 10µg of the plasmid by restriction digest using EcoRI. The esiRNA regions and vector backbones resulting from the digest were resolved by DNA agarose gel electrophoresis and were purified using QIAquick gel extraction kit (Qiagen) according to manufacturer's instructions.

The *in vitro* transcription (IVT) reactions were set up using MEGAscript T7 in vitro Transcription Kit (Ambion, Life Technologies) as follow:

Table 2.25. Reaction mixtures for IVT reactions.

10x T7 reaction buffer	1 µl
UTP (75mM)	1 µl
GTP (75mM)	1 µl
ATP (75mM)	1 µl
CTP (75mM)	1 µl
T7 enzyme mix	1 µl
PCR product / EcoRI-digested insert	4 µl
Total volume	10 µl

The IVT reaction was performed according to the following setting:

Table 2.26. Thermal cycler programme for IVT reactions.

Stages	Setting
In vitro transcription	• 12 hr at 37°C* (i.e. o/n)
Denaturation	• 3 min at 90°C • Ramp to 70°C with 0.1°C/s • 3 min at 70°C • Ramp to 50°C with 0.1°C/s • 3 min at 50°C
Annealing	• Ramp to 25°C with 0.1°C/s

0.4µl of the IVT product was mixed with 9.6µl of 2X DNA sample buffer and was run on a 1.5% DNA agarose gel, at 120V for 50 minutes. A thick band of RNA should be seen for successful IVT reactions. The rest of the IVT product was used for the subsequent digestion step.

2.4.4 Double-stranded RNA digestion

Prior to dsRNA digestion using purified recombinant GST-RNAseIII, a digestion buffer was prepared according to the recipe shown in Table 2.33. pH of the solution was adjusted to 7.9.

Table 2.27. Recipe of dsRNA digestion buffer.

Stock Solutions	Volume	[Final]
1M Tris-HCl	2 ml	20mM
5M NaCl	2.8 ml	140mM
0.1M EDTA	0.5 ml	0.5mM
1M $MgCl_2$	0.5 ml	5mM
2.5M KCl	108 µl	2.7mM
1M DTT	0.1 ml	1mM
Glycerol	5 ml	5% (v/v)
Water	89 ml	

For the digestion reaction, 9.6µl of IVT product was mixed with 86.4µl of dsRNA digestion buffer and 4µl of 1.5µg/µl GST-RNAse III. The reaction mixture was incubated on a thermomixer at 23˚C, with agitation at 1000rpm for 4 hours. Then, the temperature was increased to 37˚C and the reaction mixture was incubated with agitation at 1000rpm for another 2-3 hours. 3µl of the digestion product was mixed with 3µl 5X DNA sample buffer and was run alongside low molecular weight DNA ladder (NEB) on 2.5% DNA Agarose gel at 120V for 50 minutes. Digestion product was expected to be a smear between 18-25bp.

2.4.5 esiRNA purification

Several buffers, namely wash buffer, equilibration buffer and elution buffer, were prepared according to the recipes as summarised in Table 2.34, 2.35 and 2.36 respectively. All buffers were adjusted to pH 8.0 and kept on ice prior to use.

Table 2.28. Recipe of wash buffer.

Stock	Amount to be used	[Final]
1M Tris-HCl	2ml	20mM
5M NaCl	8ml	400mM
0.1M EDTA	1ml	1mM
Water	89ml	

Table 2.29. Recipe of equilibration buffer.

Stock	Amount to be used	[Final]
1M Tris-HCl	2ml	20mM
5M NaCl	6ml	300mM
0.1M EDTA	1ml	1mM
Water	91ml	

Table 2.30. Recipe of elution buffer.

Stock	Amount to be used	[Final]
1M Tris-HCl	2ml	20mM
5M NaCl	10.4ml	520mM
0.1M EDTA	1ml	1mM
Water	86.6ml	

200µl of Q-Sepharose was added to an empty spin column and the spin column was transferred onto a 2 ml microcentrifuge tube. 500µl of equilibration buffer was added to the spin column and the column was centrifuged at 1000g for 1 minute, and the flowthrough was discarded. This step was repeated once. All the dsRNA digestion product was then load into the column and was allowed to incubate at room temperature for 5 minutes. The spin columns were then centrifuged at 1000g for 1 minutes and the flowthrough was discarded. 500µl of wash buffer was then added, the spin columns were then centrifuged at 1000g for 1 minute and the flowthrough was discarded. The spin columns were then transferred to clean 2ml Eppendorf tubes and 300µl of elution buffer was added to each column. The spin columns together with the Eppendorf tubes were centrifuged at 1000g for 1 minute. This step was repeated once and the spin columns were discarded. Then, 300µl of isopropanol was added to the 600µl of eluate and mixed by vortexing. The mixture was then stored at -20˚C for at least 3 hours or overnight to allow precipitation of esiRNA. Following this incubation, the solution was centrifuged at 15000g for 25 minutes at 4˚C and the supernatant was discarded. The pelleted esiRNA was washed with 500µl ice-cold 70% (V/V) ethanol, centrifuged at 15000g for 15 minutes and the supernatant was discarded. This step was repeated once. The esiRNA pellet was then allowed to dry in a Speed Vac machine at 45˚C. The pellet was dissolved in 100µl RNAse-free water and concentration of esiRNA was measured using a NanoDrop Spectrophotometer ND1000 at a wavelength of 260nm.

Chapter 3 : Identification of Deubiquitylases Involved in Regulation of E-cadherin

3.1 Introduction

It has been demonstrated that the ubiquitylation of E-cadherin by the E3 ligase, Hakai can lead to its internalisation via the endocytic pathway and subsequent trafficking to the lysosome for degradation (Fujita *et al.*, 2002). However, the identity of deubiquitylases, which can reverse or regulate these processes, remains elusive. The work presented in this chapter aims to identify the deubiquitylases, which regulate E-cadherin expression level or trafficking, with the following aims: (i) to develop a robust assay to assess E-cadherin status in a human cancer cell line, (ii) to perform a siRNA DUB library screen to identify potential regulators of E-cadherin within the deubiquitylase family, (iii) to validate potential targets identified from this siRNA screen and (iv) to understand the mechanism by which the validated DUBs regulate E-cadherin.

3.2 Developing an assay to assess E-cadherin status

3.2.1 Hepatocyte Growth Factor induces scattering of epithelial cells but does not affect E-cadherin protein level

As mentioned above, ubiquitylation of E-cadherin commits it to lysosomal degradation (Fujita *et al.*, 2002; Bonazzi *et al.*, 2008) and there are multiple points during the process at which deubiquitylases can execute their regulatory roles. Hepatocyte growth factor (HGF), also known as scatter factor, is known to induce dissociation of the AJ complex (Hiscox and Jiang, 1999) and in some instance, the degradation of E-cadherin (Miura *et al.*, 2001). Stimulation of Madin-Darby Canine Kidney (MDCK) epithelial cells by HGF for example results in ubiquitylation and subsequent degradation of E-cadherin (Fujita *et al.*, 2002). In order to follow up on these observations, a panel of human cancer cell lines were stimulated with HGF and the effect on E-cadherin status was assessed. Among the tested cell lines, HGF stimulation did not affect E-cadherin in HT29 (colon cancer) and MCF7 (breast cancer) cells, whereas a slight decrease in E-cadherin was observed with DU145 (prostate cancer) and HCT116 (colon cancer) cells after 24 hours (Figure 3.1).

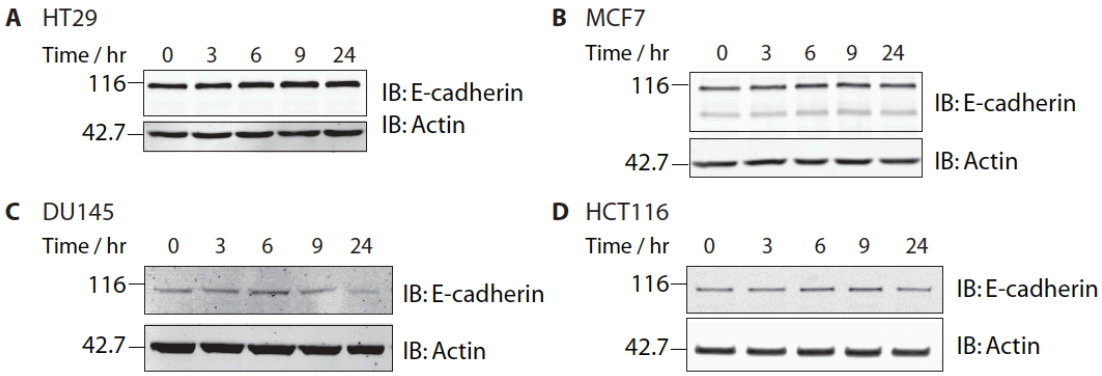

Figure 3.1. HGF stimulation did not result in a significant change in E-cadherin level. 4 different human cancer cell lines, HT29, MCF7, DU145 and HCT116 were stimulated with HGF at 50ng/ml in full serum medium for different lengths of time. Cells were lysed using NP40 lysis buffer after each time point and 20µg lysate was analysed by 10% SDS-PAGE followed by immunoblotting using HECD1 antibody against E-cadherin. Actin was used as a loading control. Experiments were repeated twice and representative blot image is presented.

In parallel to the biochemical detection of E-cadherin, the localisation of E-cadherin was examined, following HGF stimulation of the different cell lines by immunofluorescence (Figure 3.2). Unstimulated cells were tightly bound to each other, as evident by strong E-cadherin staining. The HGF stimulation resulted in the detachment of the cells, with a decrease in E-cadherin staining on the plasma membrane and a progressive change in phenotype, over the 24-hour period: Unstimulated HT29 cells were cobblestone-shaped and tightly packed with strong E-cadherin staining at cell-to-cell contacts. Following HGF stimulation, the cells slowly detached from each other and were more flattened out, and dispersing staining of E-cadherin can be seen at 24 hour time point, while membrane staining of E-cadherin was still detectable in some cells; MCF7 cells started to detach from each other following HGF stimulation, and became more flattened in shape, but still maintained partial contact at the end of the 24-hour period. The detached cells also started to form lamellipodia, which became more prominent after 9 hours; DU145 exhibited a relatively drastic change in phenotype following HGF stimulation. Unstimulated DU145 cells were rounder in shape and were attached to each other as evident by E-cadherin staining at cellular junction. Upon HGF stimulation, the cells started to come apart and at the end of 24 hours, they exhibited fibroblastic features with elongated cell shape and almost complete detachment from each other. HCT116 cells, before stimulation, were tightly packed with each other and strong E-cadherin staining was observed at cellular junctions. HGF stimulation of these cells resulted in partial detachment of cells from each other, where island of cells hollowed out and formed circles of cells, with neighbouring cells still bound to each other. In all cases, no obvious redistribution of E-cadherin into specific sub-cellular compartments was observed.

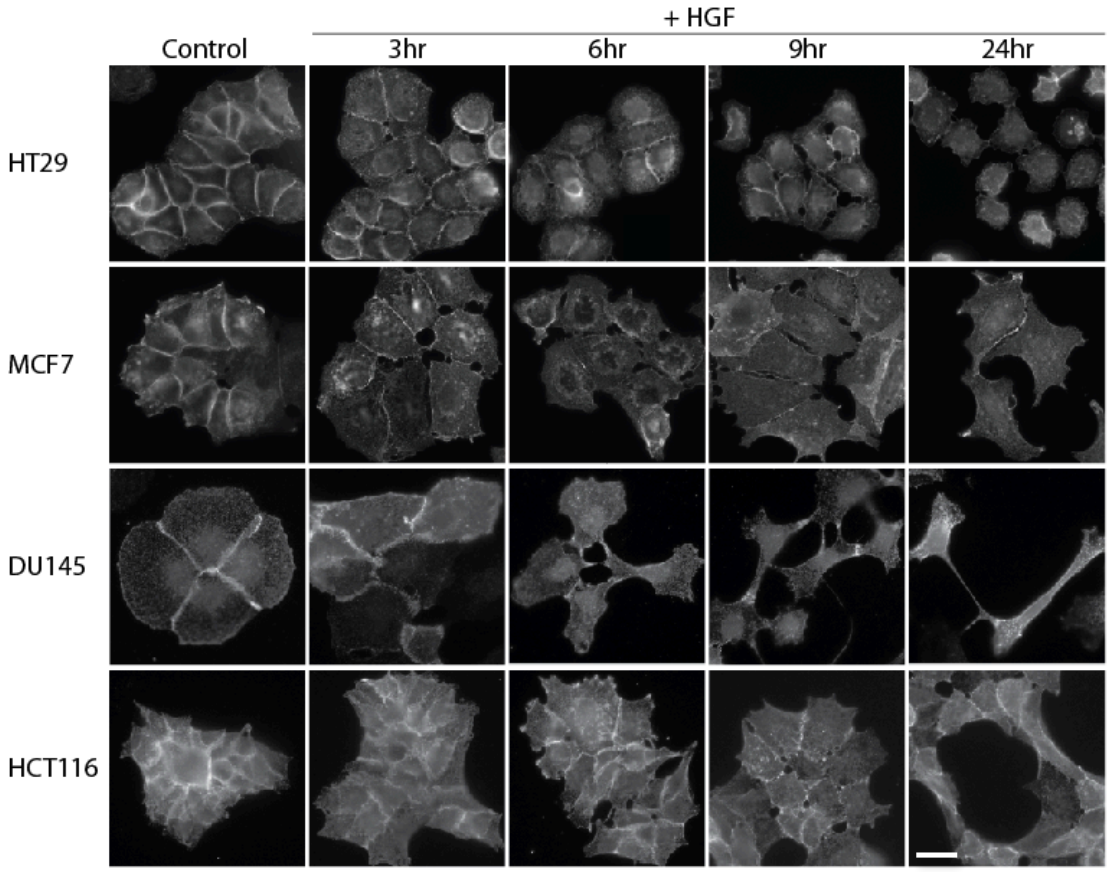

Figure 3.2. HGF stimulation resulted in dissolution of adherens junction and scattering of cancer cells. Cells were stimulated with 50ng/ml of HGF stimulation in full serum medium and were subjected to paraformaldehyde fixation at different time points. Cells were then immunostained with HECD1 antibody against E-cadherin, followed by secondary antibody. (Scale bar = 10µm).

3.2.2 E-cadherin is constitutively degraded via the lysosomal pathway in MCF7 cells

MDCK cells stimulated with HGF lose their E-cadherin as soon as 3 hours after stimulation (Fujita *et al.*, 2002). However, in the tested human cancer cell lines, despite the dissolution of adherens junction observed following HGF stimulation, loss of E-cadherin was only apparent in DU145 and HCT116 after 24 hours of stimulation. One possible explanation is that the transcriptional inactivation of E-cadherin gene expression required more than 24 hours, and the loss of E-cadherin through degradation was masked by its continual synthesis. Therefore, to check if E-cadherin was degraded following HGF stimulation, I repeated the experiments in the

presence of a proteasomal inhibitor, Lactacystin or lysosomal ATPase proton pump inhibitor, Folimycin (Figure 3.3). The use of the inhibitors was expected to lead to accumulation of full length or degradative intermediate products of E-cadherin. Interestingly, in MCF7 cells, Folimycin but not Lactacystin treatment revealed a lower molecular weight band of E-cadherin, which accumulated in a time-dependent manner (Figure 3.3A). The E-cadherin fragment ran at around 80kDa, corresponding to the extracellular domain of E-cadherin. Notably, the accumulation of this E-cadherin fragment was independent of HGF stimulation, suggesting constitutive turnover of E-cadherin via the lysosomal pathway. When cells were treated with a combination of Lactacystin and Folimycin, the accumulation of this E-cadherin fragment was less pronounced as compared to cells treated with Folimycin alone. This could be explained by the fact that the turnover of membrane protein requires the cooperation of proteasomal and lysosomal degradation pathway, as was previously shown for Met receptor (Hammond *et al.*, 2001). No obvious changes in E-cadherin level were observed with different combinations of treatment in DU145 cells (Figure 3B).

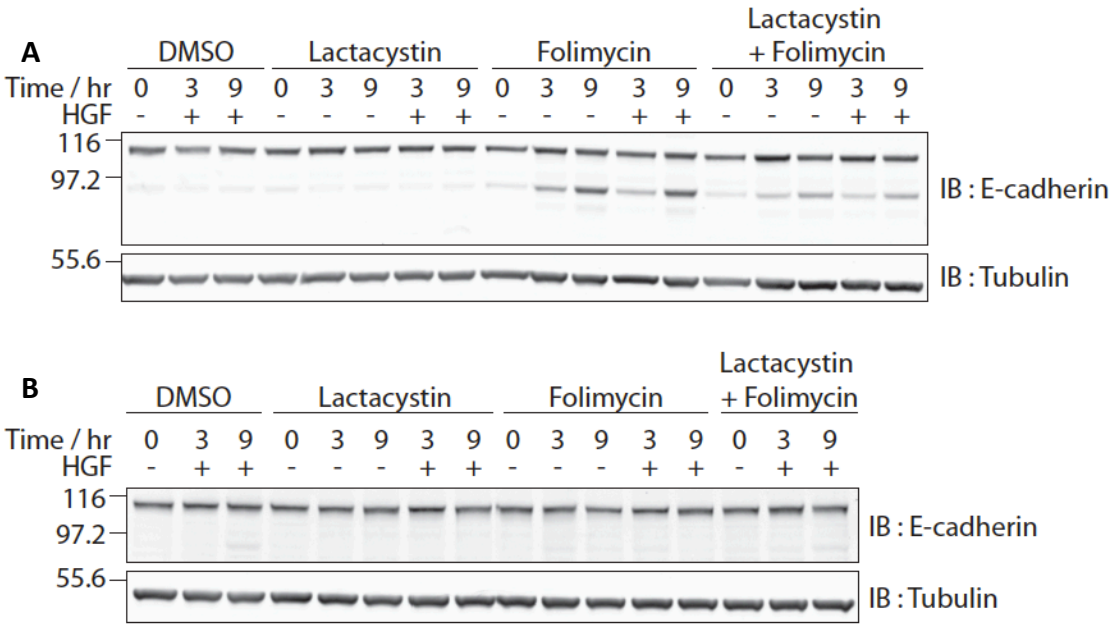

Figure 3.3. E-cadherin is constitutively degraded in MCF7, but not DU145, cells via the lysosomal pathway. (A) MCF7 and (B) DU145 cells were incubated ± 50ng/ml of HGF in the presence or absence of 10μM Lactacystin and/or 100nM of Folimycin, for indicated lengths of time. Cells were lysed and analysed by SDS-PAGE, followed by immunoblotting for E-cadherin. Western Blot images were acquired using an infra-red scanner (Odyssey, LICOR).

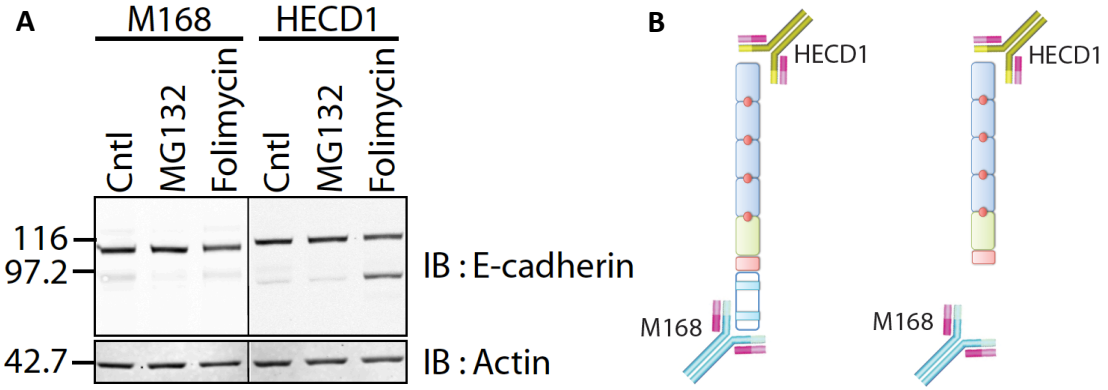

Figure 3.4. The 80kDa E-cadherin fragment corresponds to the extracellular domain. (A) MCF7 cells were treated with 10µM MG132 or 100nM Folimycin for 9 hours before lysis. Cell lysates were analysed by SDS-PAGE followed by transfer to nitrocellulose membrane. The blot was probed with M168 or HECD1 primary antibodies, and images were acquired using an infra-red scanner (Odyssey, LICOR). **(B)** Topology of E-cadherin antibody binding sites. The HECD1 antibody binds to extracellular domain of E-cadherin whereas the M168 antibody binds to the cytoplasmic domain of E-cadherin.

To characterise the 80kDa E-cadherin fragment, two different E-cadherin antibodies were used, namely M168, which recognises the cytoplasmic C-terminus and HECD1, which recognises the extracellular N-terminus. As seen in Figure 3.4, the 80kDa E-cadherin fragment was only recognised by the HECD1 antibody but not M168, indicating it was the extracellular domain of E-cadherin.

Observation by immunofluorescence miscroscopy of MCF7 cells treated with Folimycin showed focal accumulation of the E-cadherin staining, in the perinuclear area and inhibited full scattering of MCF7 cells following HGF stimulation (Figure 3.5). To determine the subcellular compartment which the E-cadherin fragment occupies, MCF7 cells were labelled with several cellular compartment markers, namely EEA1 for the early endosomes (Mu *et al.*, 1995) (Figure 3.6), CD63 (Metzelaar *et al.*, 1991) (Figure 3.7) and LAMP1 (Carlsson *et al.*, 1988) (Figure 3.8) for lysosomes, and colocalisation of E-cadherin with these markers were determined using confocal microscopy. The E-cadherin fragment accumulated following Folimycin treatment was mostly excluded from the early endosomes, as no colocalisation was observed between E-cadherin and EEA1. Noticeably, the Folimycin treatment resulted in

change in distribution of the EEA1 markers: punctae of EEA1 positive compartments were observed throughout the cytoplasm in non-treated cells, and at least after 6 hours of Folimycin treatment, the number of EEA1 positive compartments was significantly reduced.

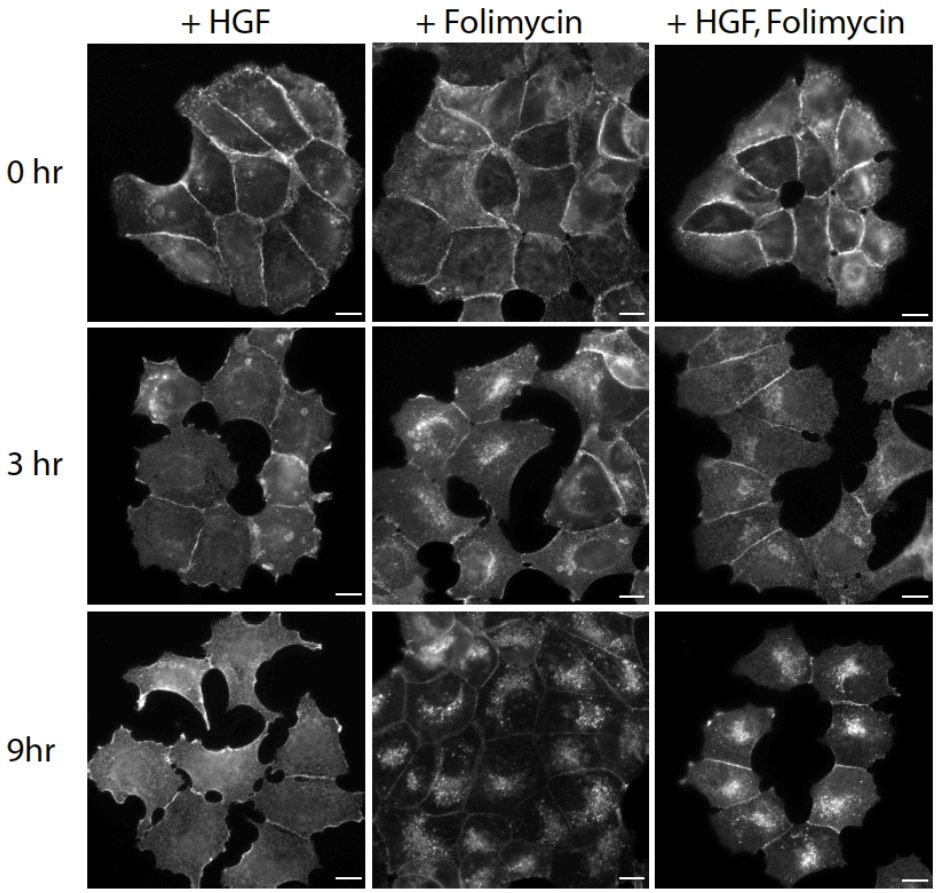

Figure 3.5. E-cadherin accumulated in the perinuclear region after Folimycin treatment. MCF7 cells were treated with 50ng/ml HGF and/or 100nM Folimycin for indicated lengths of time. Cells were then fixed with 4% paraformaldehyde and immunostained with HECD1 antibody. Scale bar = 10µm.

E-cadherin showed colocalisation with both the lysosomal markers, CD63 and LAMP1, although not all E-cadherin positive compartments were co-labelled with either of the markers. It is also worth noting that Folimycin treatment resulted in a change in distribution of the lysosomal marker positive compartment: whereas untreated cells showed diffuse cytoplasmic staining of CD63 and LAMP1, lysosomal marker positive compartments exhibited peri-nuclear accumulation in Folimycin

treated cells. Together, the results in this section suggest that in MCF7 cells, E-cadherin is constitutively internalised and is degraded via the lysosomal pathway.

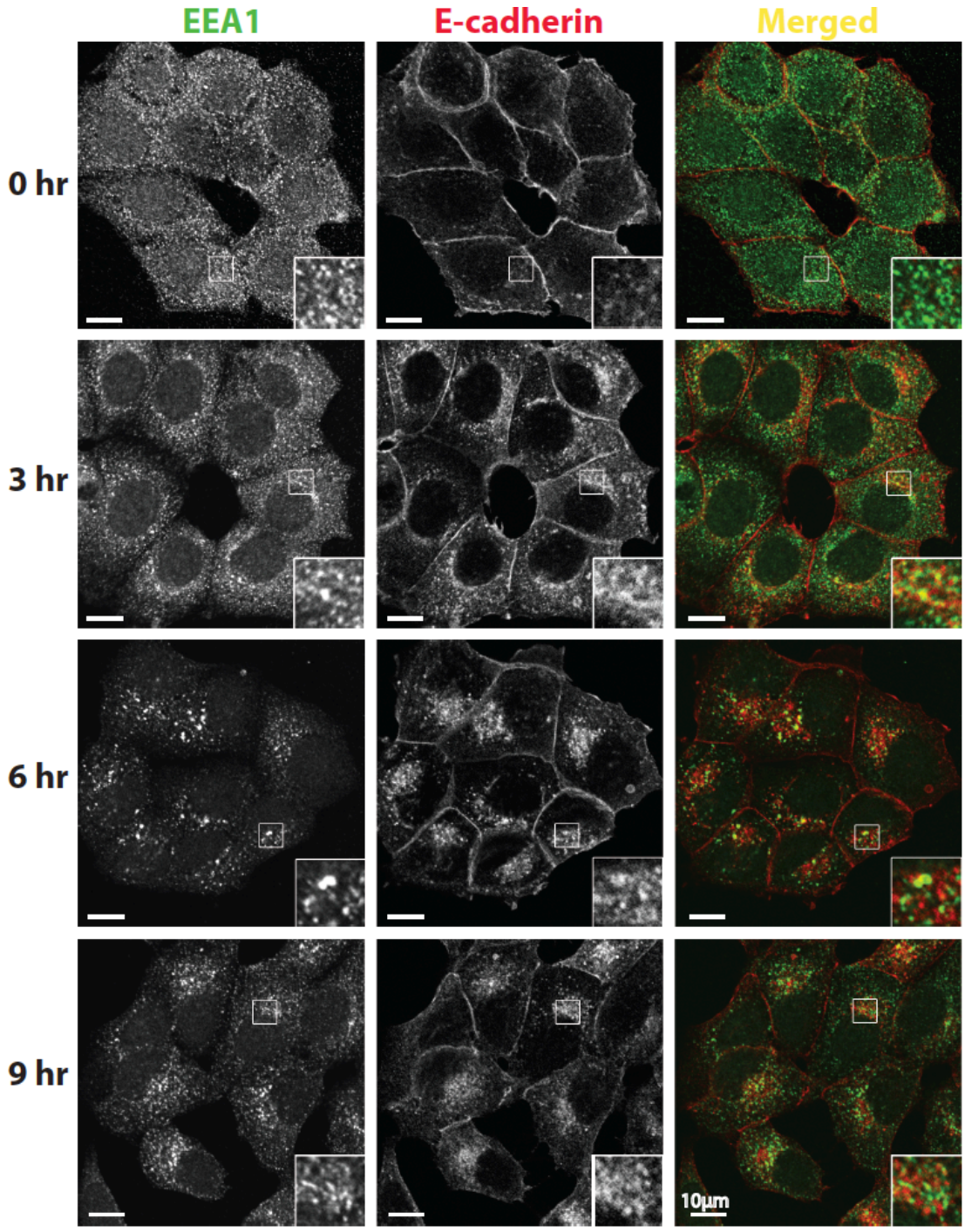

Figure 3.6 E-cadherin fragment accumulated following Folimycin treatment does not colocalise with an early endosomal marker. MCF7 cells were treated with 100nM Folimycin for indicated lengths of time. Cells were then fixed with 4% paraformaldehyde and immunostained with HECD1 and EEA1 antibodies. Images were acquired using a confocal microscope and laser intensity was adjusted for each image to avoid oversaturation of signal.

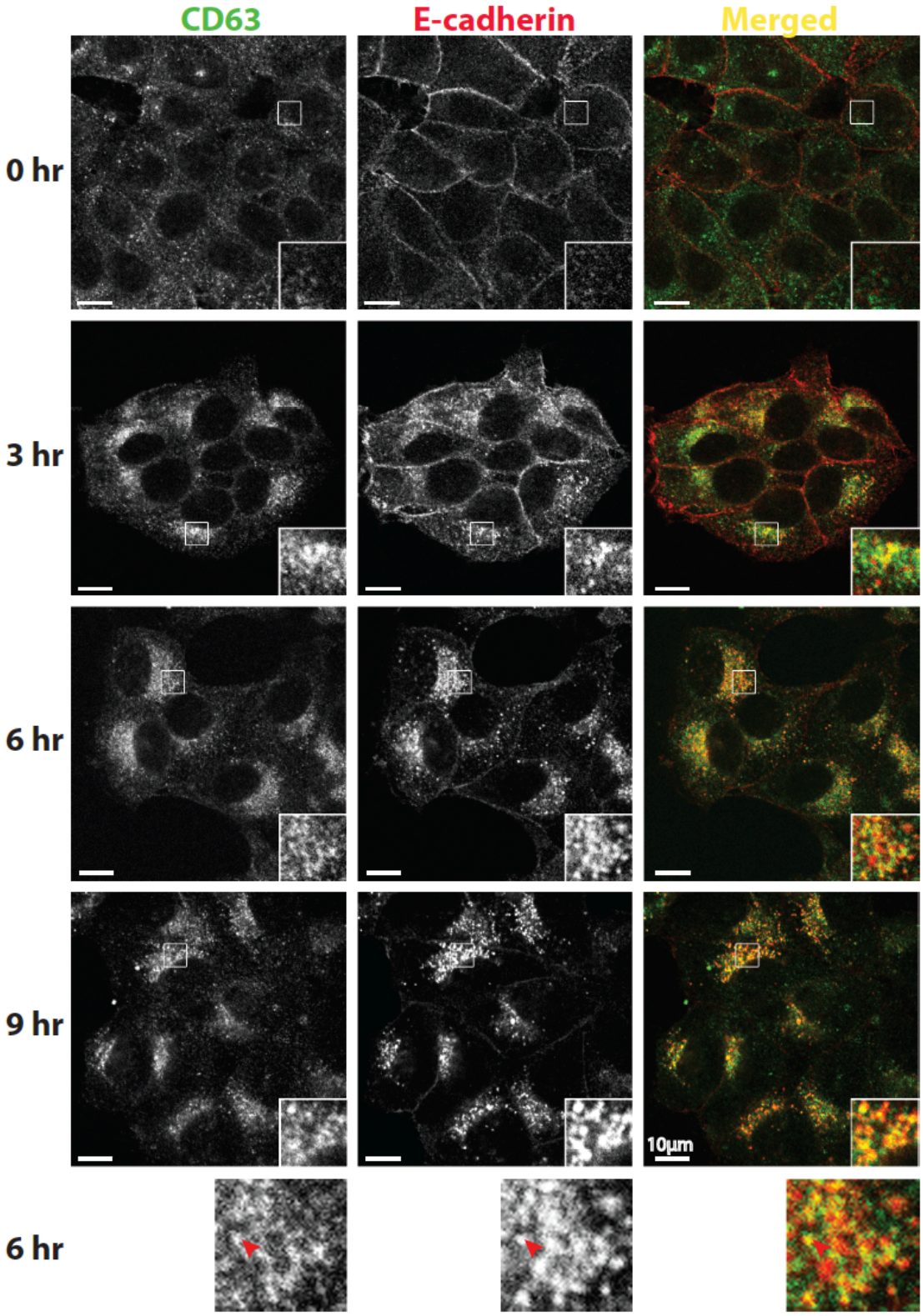

Figure 3.7 E-cadherin fragment accumulated following Folimycin treatment colocalises with CD63. MCF7 cells were treated with 100nM Folimycin for indicated lengths of time. Cells were then fixed with 4% paraformaldehyde and immunostained with EP700Y antibody against E-cadherin and anti-CD63 antibody. Images were acquired using a confocal microscope and laser intensity was adjusted for each image to avoid oversaturation of signal.

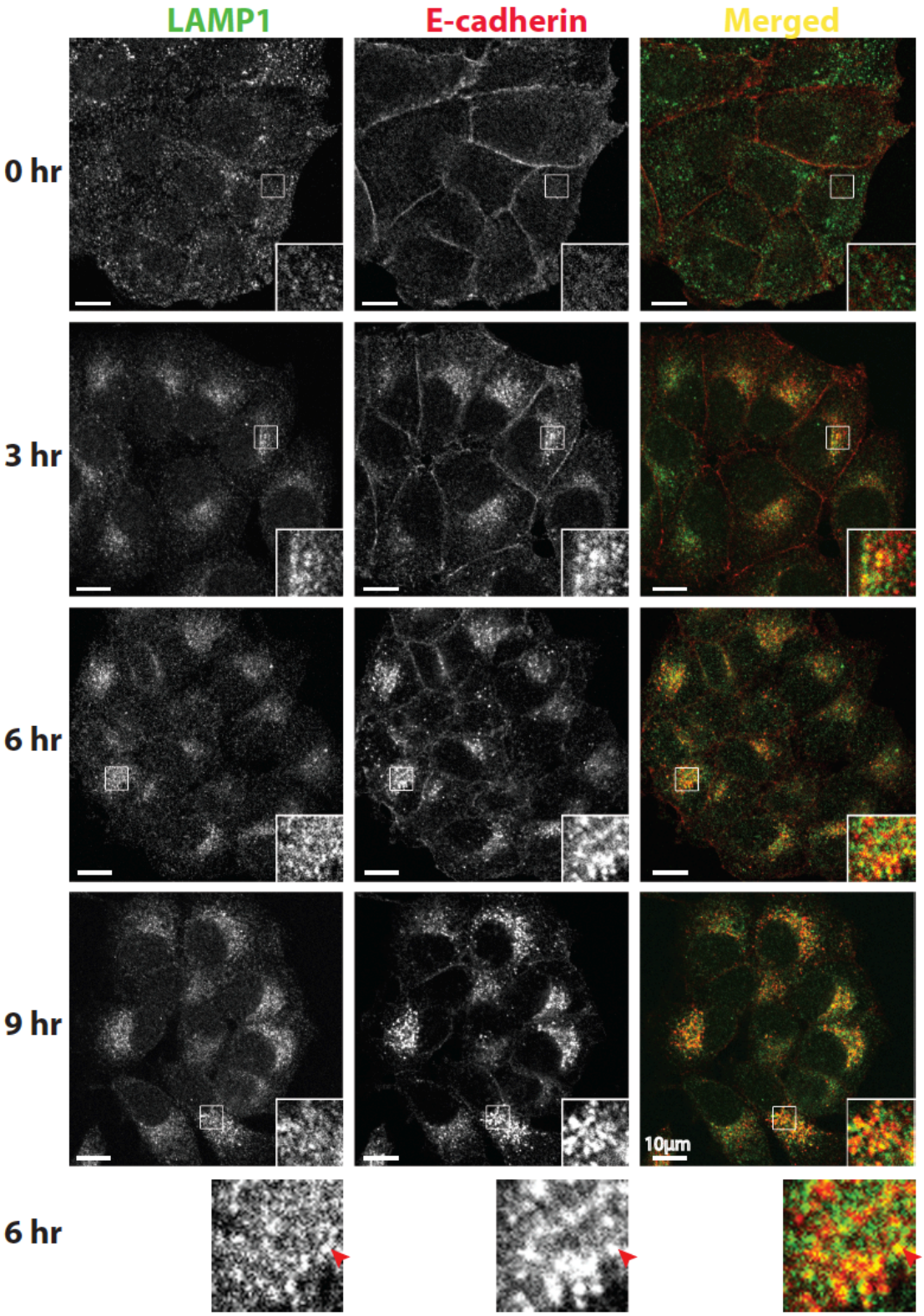

Figure 3.8 E-cadherin fragment accumulated following Folimycin treatment colocalises with LAMP1. MCF7 cells were treated with 100nM Folimycin for indicated lengths of time. Cells were then fixed with 4% paraformaldehyde and immunostained with EP700Y antibody against E-cadherin and anti-LAMP1 antibody. Images were acquired using a confocal microscope and laser intensity was adjusted for each image to avoid oversaturation of signal.

3.2.3 Identification of DUBs regulating E-cadherin in MCF7

As the accumulation of 80kDa E-cadherin following Folimycin treatment was a very robust effect and could possibly reflect a ubiquitin dependent sorting event (Raiborg and Stenmark, 2009), it was used as a biochemical read out for a siRNA DUB library screen. The ratio of the full length E-cadherin to 80kDa fragment should serve as an indicator of the trafficking status of E-cadherin, where a high ratio indicates stabilisation of E-cadherin and/or less degradation, while a low ratio indicates degradation of the molecule is favored. Another parameter I monitored was the total E-cadherin level, which was represented by the sum of full length E-cadherin and the 80kDa fragment. This could identify DUBs involved in transcription or translation as well as protein turnover. To carry out the screen, siRNA knockdown of a panel of 92 human DUBs was performed in MCF7 cells, using a DUB siRNA library purchased from QIAGEN (See section 2.2.3 for experimental procedures). Cells were allowed to grow for 72 hours following which they were treated with Folimycin for 6 hours at 100nM. After that, cells were lysed and the lysates were analysed by SDS-PAGE in combination with immunoblotting. The experimental workflow is shown in Figure 3.9.

I have run 3 sets of gels, i.e. 3 technical repeats of the screen samples. The technical repeats were essential as western blotting is quite a variable technique, in that the transfer of proteins from polyacrylamide gel to nitrocellulose membrane is not always consistent and hence results based on one gel only may be biased (Aksamitiene *et al.*, 2007). Figure 3.10 shows a representative blot image from the screen. At first glance, as compared to siControl sample, knockdown of DUBs such as USP14, USP19 and USP41 resulted in a stronger accumulation of the 80kDa E-cadherin fragment, whereas the silencing of USP44 and OTUB2 resulted in less accumulation of the 80kDa fragment (highlighted in bold in Figure 3.10). Besides, a total decrease in E-cadherin level was observed with the knockdown of USP32 and USP38.

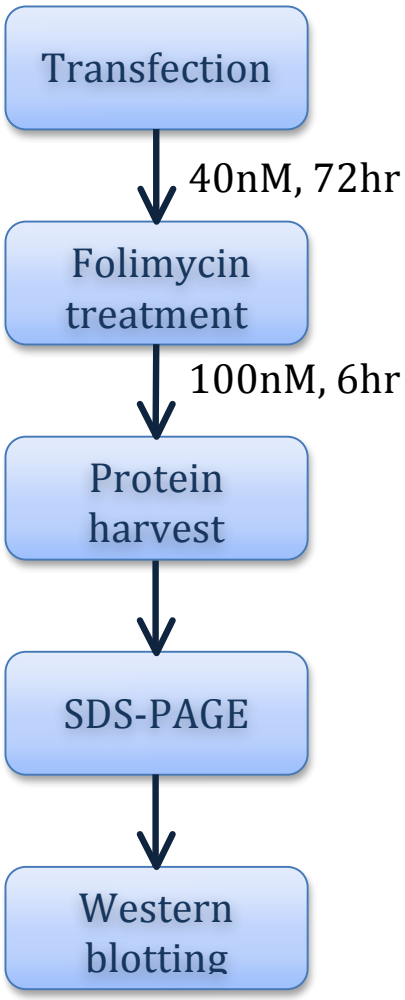

Figure 3.9. Experimental workflow of siRNA DUB library screen.

Densitometric analysis was performed to determine relative amounts of proteins that were present and statistical analysis was done in several ways. To identify DUBs whose knockdown affect E-cadherin trafficking, the ratio of the full length E-cadherin to the 80kDa E-cadherin fragment was first derived. Then, all samples were ranked according to the deviation of their ratios from that of the non-targeting control on corresponding gels (Figure 3.11A). As an alternative approach, the samples were ranked based on the deviation of their ratios from the median of the combined (4 gels) data set (Figure 3.11B) and the median sample on each gel (Figure 3.11C). The different statistical approaches provided similar top hits, although in slightly different order. USP40, USP38, OTUB2 and USP8 were the top hits, which when silenced, resulted in increase in the full length E-cadherin to 80kDa

fragment ratio, as determined by all 3 statistical approaches. This observed increase in ratio was due to less accumulation of the 80kDa E-cadherin fragment (Figure 3.10).

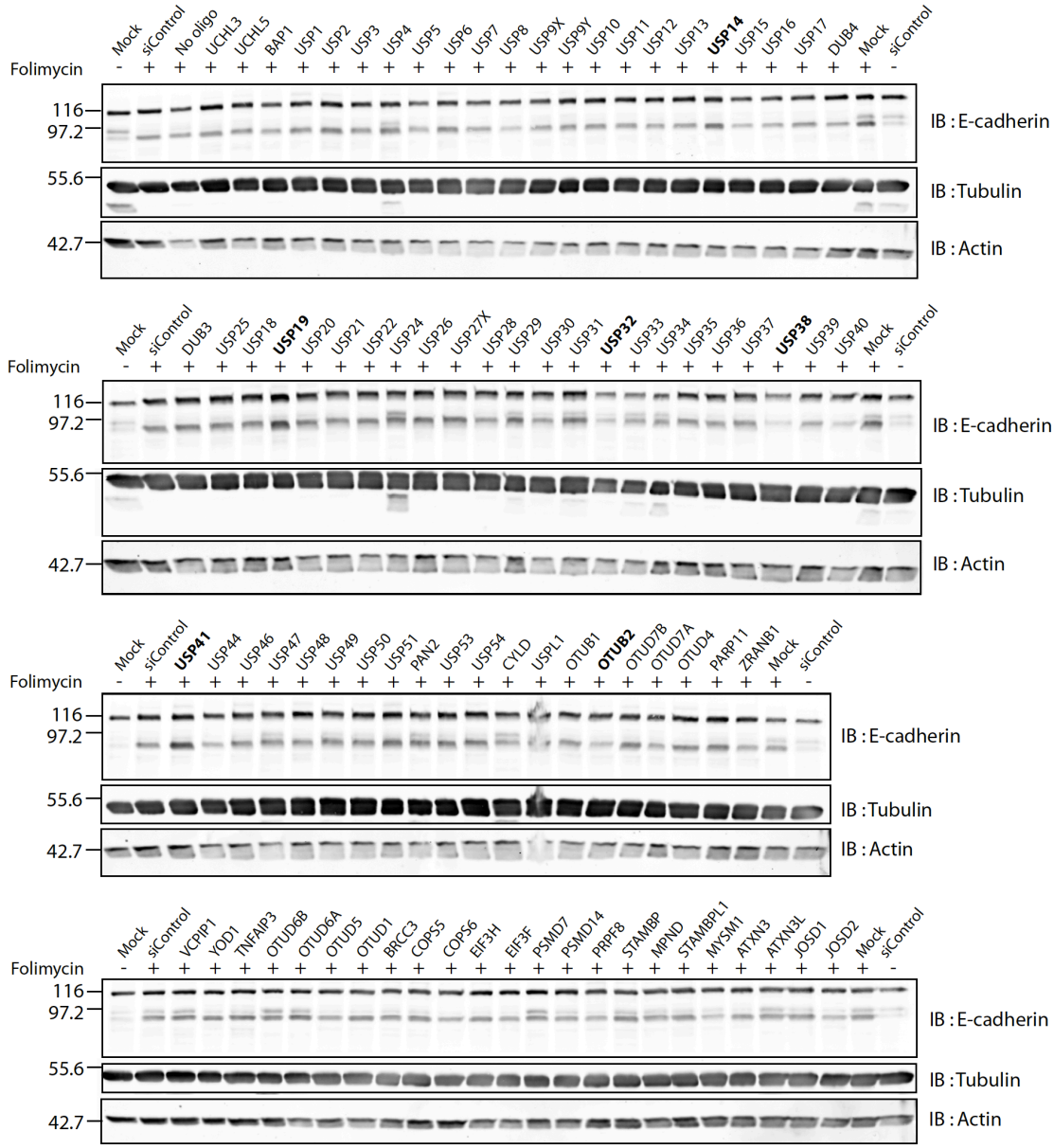

Figure 3.10. siRNA DUB library screen to identify DUBs regulating E-cadherin in MCF7. MCF7 cells were transfected with pools of 4 single oligos, each pool targeting 1 of 92 DUBs, at a final concentration of 40nM. Cells were treated with 100nM Folimycin for 6 hours 72 hours post-transfection. Cells were lysed using NP40 lysis buffer. Lysates were run on 0.8% SDS-PAGE gels and transferred to nitrocellulose membrane before immunoblotting with HECD1 antibody for E-cadherin. Tubulin and actin were used as loading control. Gel images were acquired by an infra-red scanner (Odyssey, LICOR).

USP41, USP4, USP14, USP6, USP17 and OTUD6A, on the other hand, were the top hits, which when depleted, resulted in a decrease in the ratio. Among these, it was quite apparent, judging from the gel image, that the decrease in ratio observed with knockdown of USP41 and USP14, was due to a more pronounced accumulation of the 80kDa E-cadherin fragment.

Candidate DUBs, which appeared in the top 3 as identified by at least 2 out of the 3 different quantitation methods, were then selected for subsequent deconvolution experiments to evaluate the possibility of off-target effects (Figure 3.12). STAMBP and STAMBPL1 (McCullough *et al.*, 2004; Nakamura *et al.*, 2006), although was only identified by 1 statistical approaches as top hit, were also included as they are endosomal-associated DUBs and may therefore affect E-cadherin trafficking. The observed increase in full length E-cadherin to 80kDa E-cadherin fragment ratio following knockdown of USP12 and USP40 was likely due to an off-target effect as only 1 out of the 4 oligos recapitulated the pool knockdown effect, where the ratio was increased. For all the other DUBs, at least 3 out of 4 oligos reproduced the pool knockdown effect, indicating a genuine effect on the profile of E-cadherin as a result of the loss of the corresponding DUBs. Among these, the change in ratio due to knockdown of USP38 was the most dramatic, and 3 out of 4 single oligos had the same effect as the pool oligos.

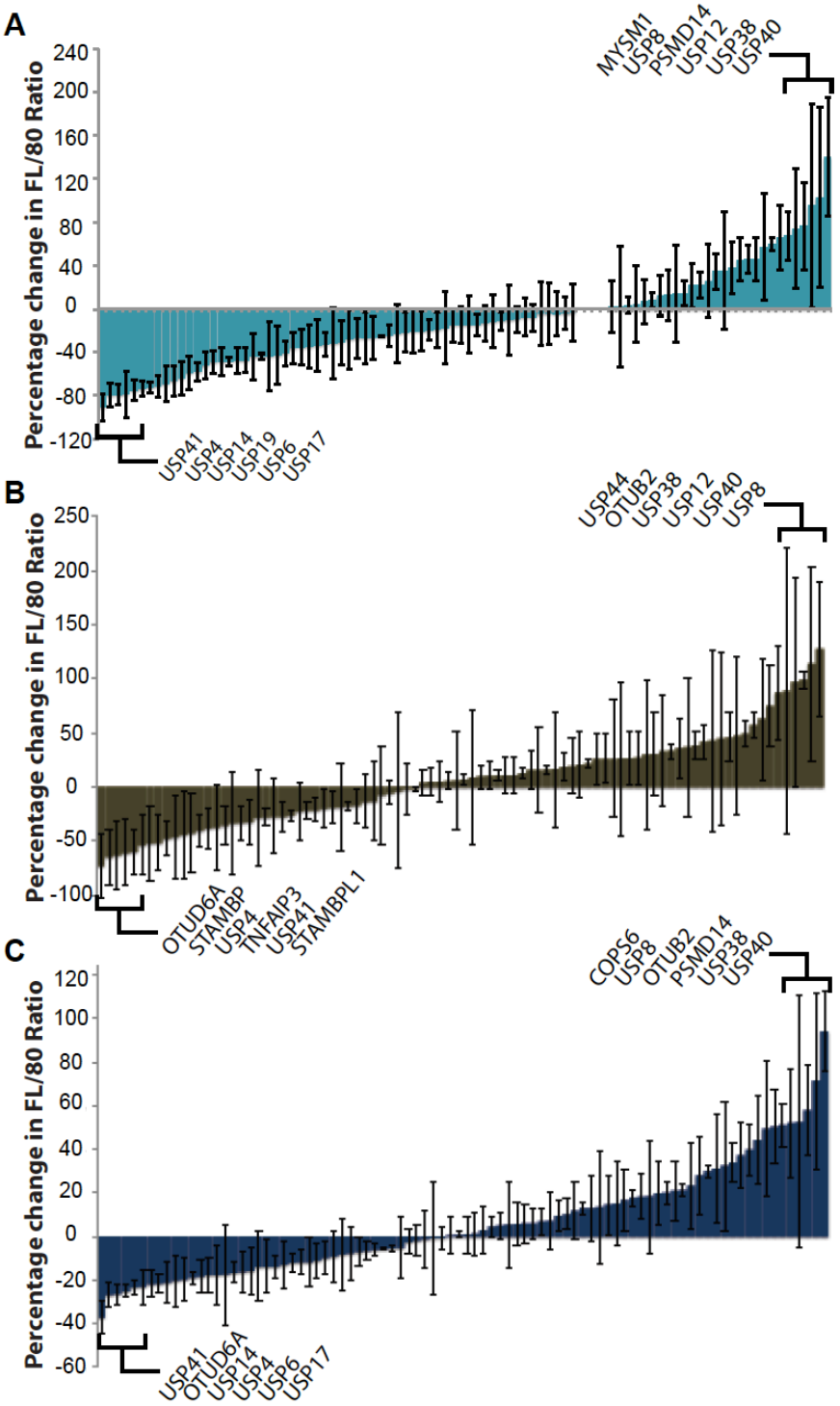

Figure 3.11. Change of full length E-cadherin to 80kDa E-cadherin fragment ratio following DUBs knockdown. Densitometric analysis was performed using ImageJ to determine relative amount of full length E-cadherin and 80kDa E-cadherin fragment for each knockdown sample, and ratios between the two were derived. The ratio of each sample was then normalised and expressed as a percentage change relative to the **(A)** siControl sample on corresponding gel and **(B)** median of the whole data set, i.e. ratios of samples from all gels and **(C)** median sample on corresponding gel. Data shown represents average of 3 technical repeats and only the top 6 candidates on both positive and negative ends were shown. Error bars indicate standard deviation of data.

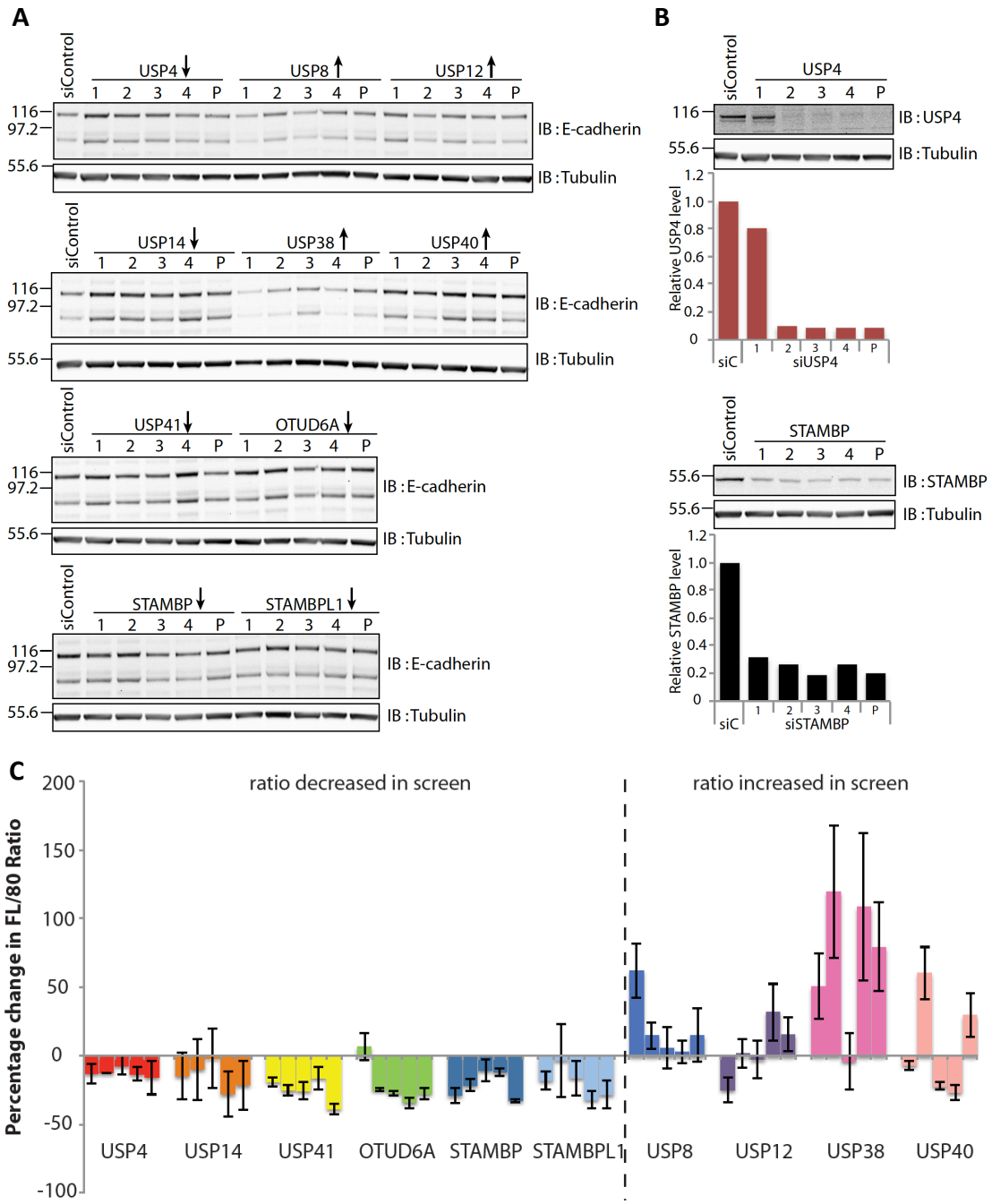

Figure 3.12. **Deconvolution of targets which altered full length E-cadherin to 80kDa E-cadherin fragment ratio. (A)** and **(B)** MCF7 cells were transfected with single oligos against candidate DUBs at 40nM for 72 hours. Cells were then treated with 100nM Folimycin for 6 hours. Cell were lysed using NP40 lysis buffer. Lysates were resolved by SDS-PAGE and immunoblotted for E-cadherin, USP4 and STAMBP. Tubulin was used as loading control. Gel images were acquired by an infra-red scanner (Odyssey, LICOR). Arrows indicate direction of change of E-cadherin level following knockdown of the chosen DUBs in screen. **(C)** Densitometric analysis was performed using Image J to determine band intensity and change in full length E-cadherin to 80kDa fragment ratio with respect to siControl sample was determined. Data shown represents average of 3 technical repeats, and the 5 columns for each DUB represents oligos 1, 2, 3, 4 and pool oligos respectively. Error bars indicate standard deviation of data.

I have also determined the change in total E-cadherin level, i.e. the summed intensity of both full length E-cadherin and 80kDa E-cadherin fragment, normalised to tubulin or actin, following knockdown of individual DUBs (Figure 3.13). The top candidate DUBs, which when depleted resulted in significant reduction in E-cadherin level, were USP38, USP32, MYSM1, BAP1 and USP16, whereas the DUBs for which depletion resulted in an increase in total E-cadherin level were USP19, USP22, USP6, OTUD1, USP20, and USP31. Out of the 12 top candidates shown for data normalised to tubulin or actin, only 6 (USP19, USP31, OTUD1 on positive end and USP38, USP32 and USP16 on negative end) were identified with both methods of normalisation but in different orders. Among the listed candidates, the identification of USP19, USP32 and USP38 was anticipated based on visual judgment of the immunoblots, where more intense bands of both full length and truncated E-cadherin were observed for USP19, and the bands for USP32 and USP38 were relatively faint. While a significant increase in total E-cadherin level was visible for USP41 knockdown sample (Figure 3.10), the DUB was however not listed as 1 of the top hits. This highlights the reliability issue of the quantitation results due to normalisation to tubulin or actin, since the bands for these two loading controls were not sharp.

Based on the quantitation of the change in full length E-cadherin level, USP19, USP32, USP38 and USP41 were chosen for deconvolution (Figure 3.13), where the latter two DUBs were already chosen based on the effect of their depletion on E-cadherin ratio (Figure 3.12). For USP32, although 3 out of 4 oligos resulted in depletion of E-cadherin level, the effect was less marked compared with the original screen (<20% v.s. ~40%). For USP38, oligos 1, 2 and 4 resulted in >40% depletion of E-cadherin, reproducing the pool knockdown effect observed in the screen. For both USP19 and USP41, 3 out of 4 oligos resulted in a minimum increase of 40% in total E-cadherin level.

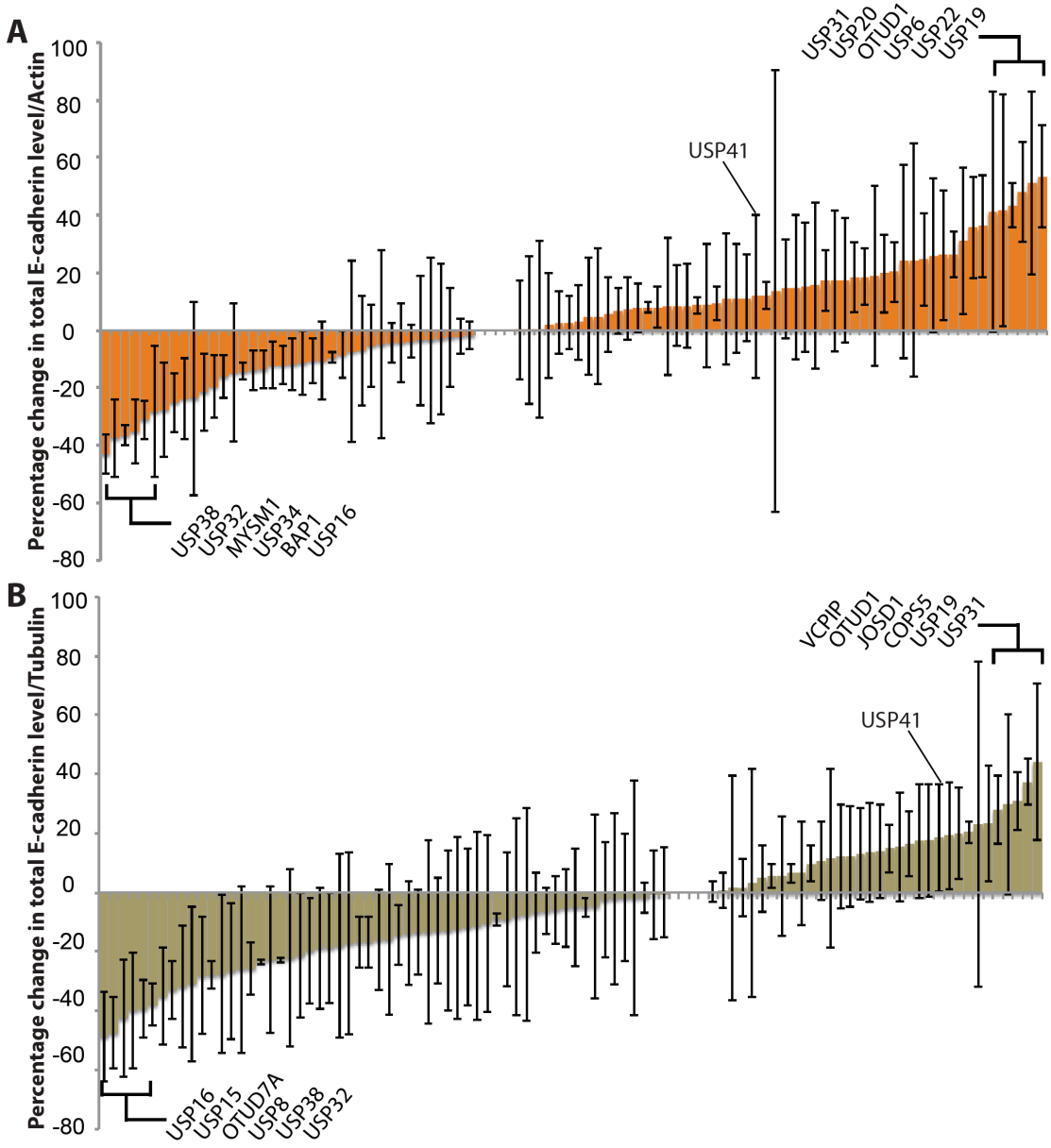

Figure 3.13. Change in total E-cadherin level following knockdown of individual DUBs. Densitometric analysis was performed using ImageJ to determine relative amount of full length E-cadherin and 80kDa E-cadherin fragment for each knockdown sample, and the sum of the two was calculated and was normalised to **(A)** actin and **(B)** tubulin. The normalised E-cadherin levels were then expressed as a percentage deviation from the siControl samples. Samples were then ranked in ascending order. Data shown represents average of 3 technical repeats and only the top 6 candidates on both positive and negative ends were shown. Error bars indicate standard deviation of each sample.

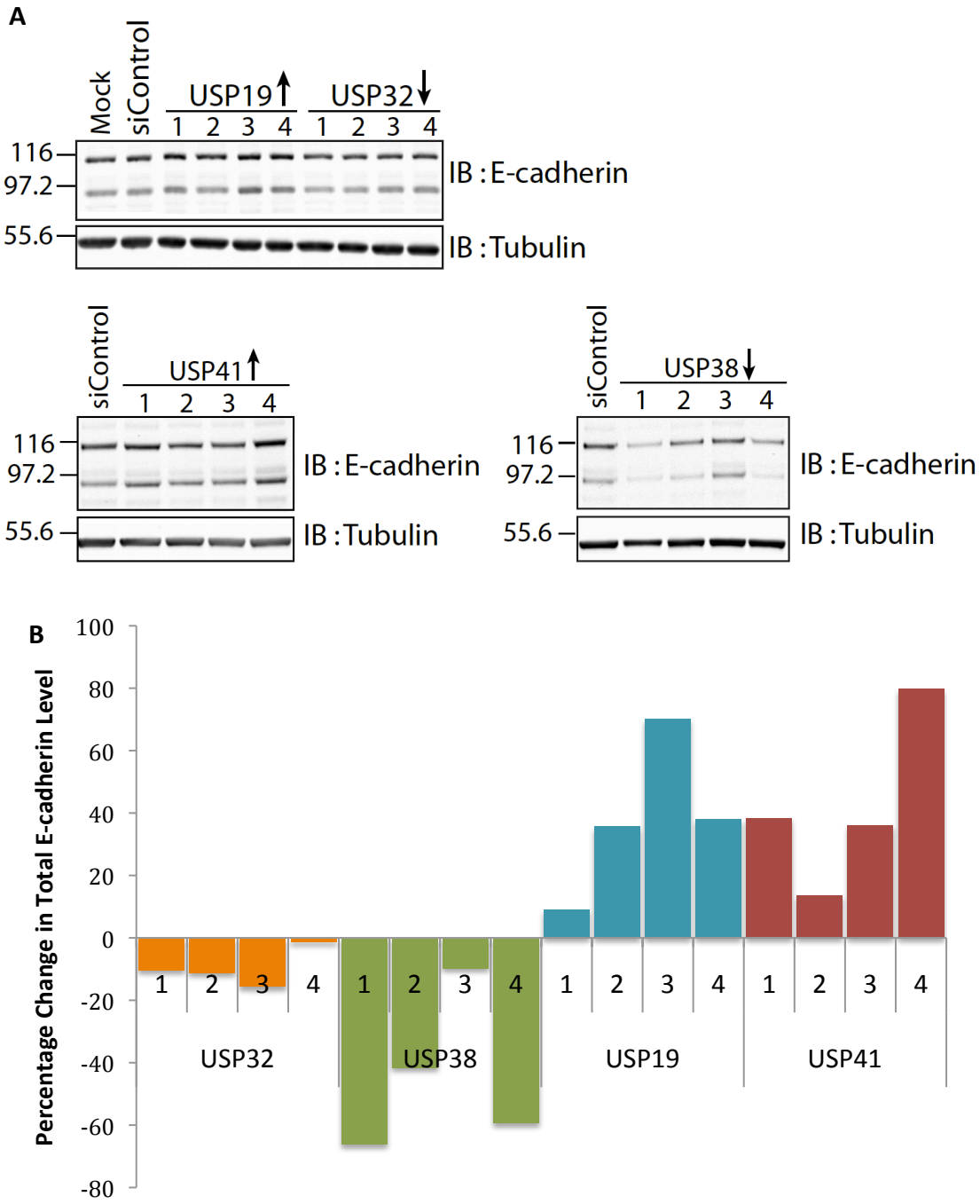

Figure 3.14. Deconvolution of targets which altered total E-cadherin level. (A) MCF7 cells were transfected with single oligos against candidate DUBs at a final concentration of 40nM. Cells were treated with 100nM Folimycin for 6 hours, 72 hours post-transfection. Cells were lysed using NP40 lysis buffer. Lysates were resolved by SDS-PAGE and immunoblotted for E-cadherin and tubulin. Gel images were acquired by infra-red scanner (Odyssey, LICOR). **(B)** Densitometric analysis was performed using Image J software to determine band intensity and sum of full length E-cadherin and 80kDa E-cadherin fragment was calculated, and normalised to tubulin level. Change in total E-cadherin level from that of siControl was then determined and represented in graph. Data shown represents average of 2 technical repeats.

3.2.4 Identification of DUBs regulating E-cadherin in A549

Two colleagues in our laboratory, Joseph Sacco and Han Liu, performed a large scale DUB library screen in A549 cells and the NP40 cell lysates were processed for SDS-PAGE and aliquoted into 96-well plates, serving as a resource for the identification of DUBs playing roles in regulating a cellular process or stability of a protein in steady state conditions (Figure 3.15). I have utilised this resource as a parallel mean to identify DUBs involved in E-cadherin regulation.

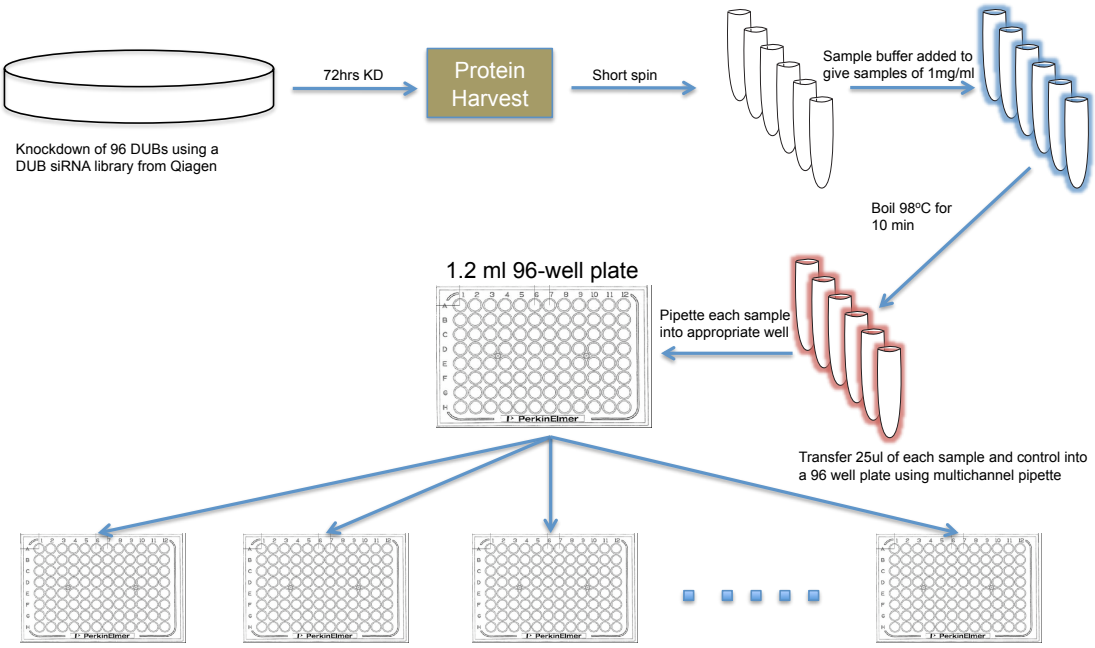

Figure 3.15. Large scale DUB siRNA library screen. A549 cells were seeded in 15cm dish and cells were either mock-transfected or transfected with non-targeting siRNAs or pool of 4 single oligos against 1 of 92 DUBs, at a final concentration of 40nM. Cells were lysed using NP40 lysis buffer 72 hours post-transfection and following protein assay, lysates were boiled with sample buffer and adjusted to a concentration 0.8µg/µl. Finally, 25µl of each knockdown samples was arrayed on 96-well plates. (experiment was performed by Liu and Sacco)

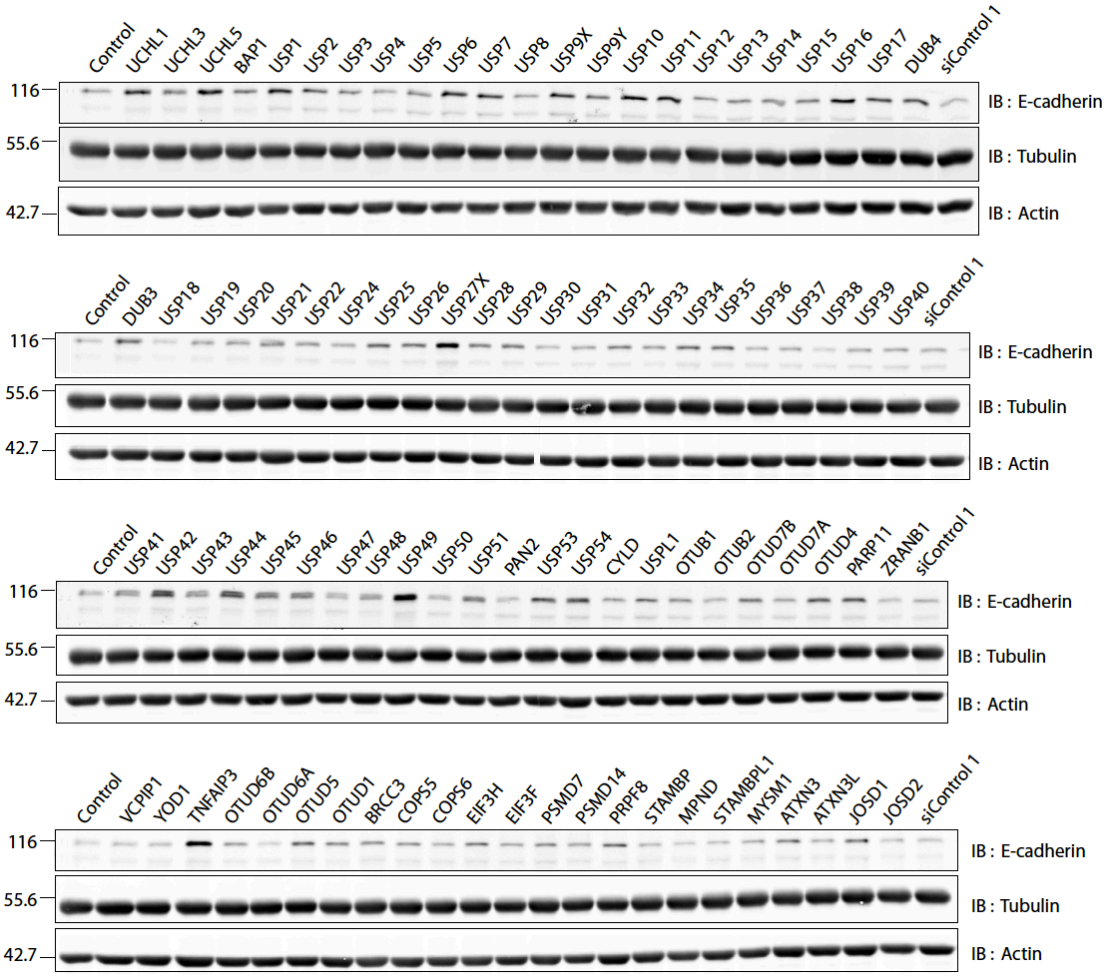

Figure 3.16. siRNA DUB library screen to identify DUBs regulating E-cadherin in A549. NP40 lysates from large scale DUB screen performed by Sacco and Liu were resolved on 10% SDS-PAGE gels, followed by Western Blotting. Blots were probed with antibodies against E-cadherin, tubulin and actin. Gel images were acquired by an infra-red scanner (Odyssey, LICOR). Notes : Control = mock knockdown without oligo; siControl1 = knockdown with non-targeting oligos.

Figure 3.16 shows the representative blot image acquired for this set of experiments. At first glance, depletion of a number of DUBs on Gel1 resulted in visible increase in E-cadherin level as compared to the controls. Across all the gels, TNFAIP3, USP27X and USP49 resulted in the most dramatic increase in E-cadherin level.

Densitometric analysis of the blot images was performed to give a quantitative measurement of the E-cadherin level (Figure 3.17). The top DUB candidates, which when silenced resulted in significant increase in E-cadherin level,

as determined by both quantitation approaches, were TNFAIP3, USP27X, USP49 and USP42, in agreement with the visual observation. On the other end of the graph were ZRANB1, OTUB2, USP38 and PAN2, which when silenced resulted in depletion of E-cadherin. OTUD6A did not appear in the top hit lists, although it was quite close to the negative end.

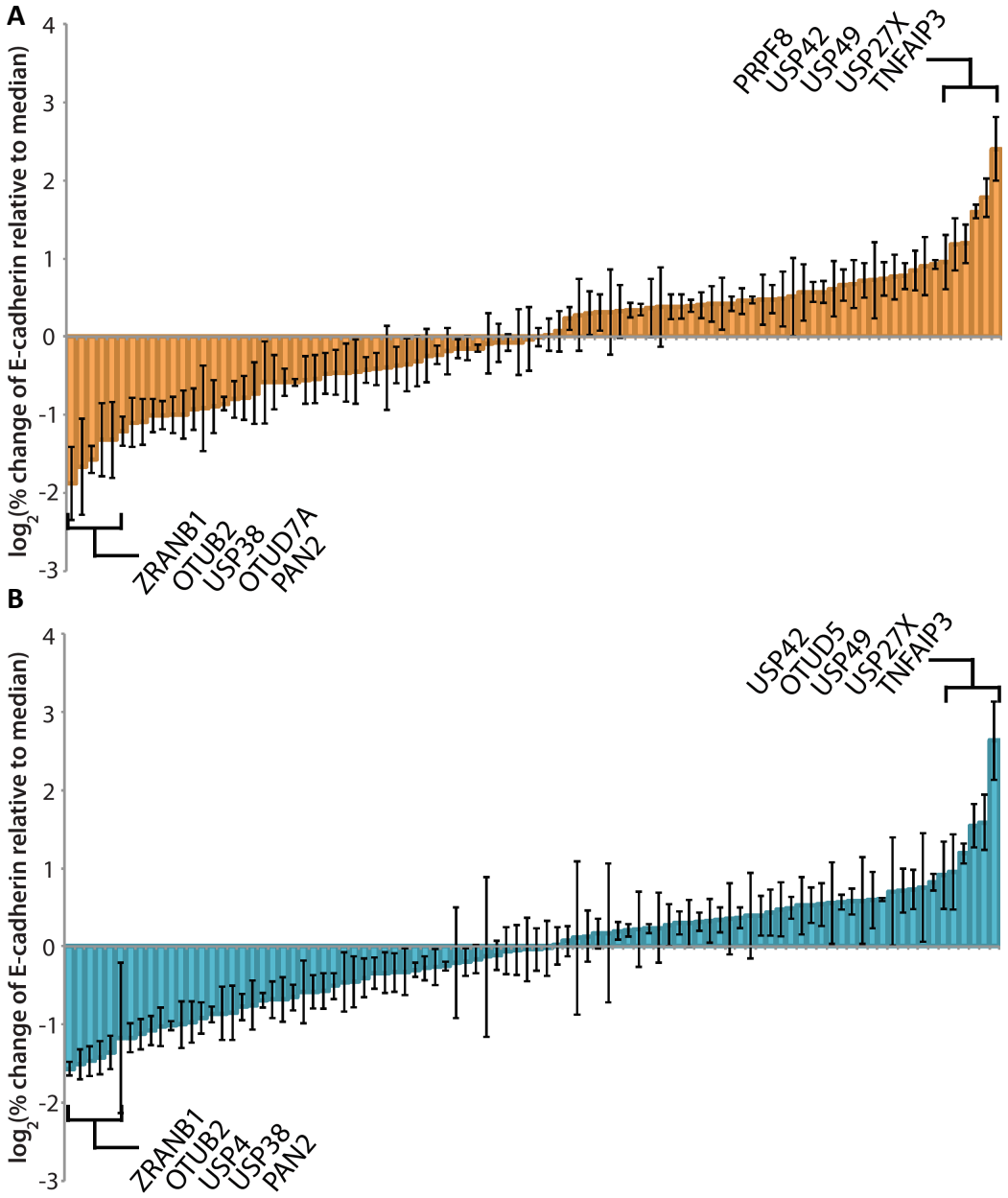

Figure 3.17. Change in E-cadherin level following knockdown of individual DUBs. Densitometric analysis was performed using ImageJ to determine relative amount of E-cadherin for each knockdown sample and E-cadherin level was normalised to **(A)** tubulin or **(B)** actin. The percentage change in E-cadherin level relative to that of the median sample was calculated and log-transformed. Data shown represents average of 3 technical repeats and error bars represent standard deviation.

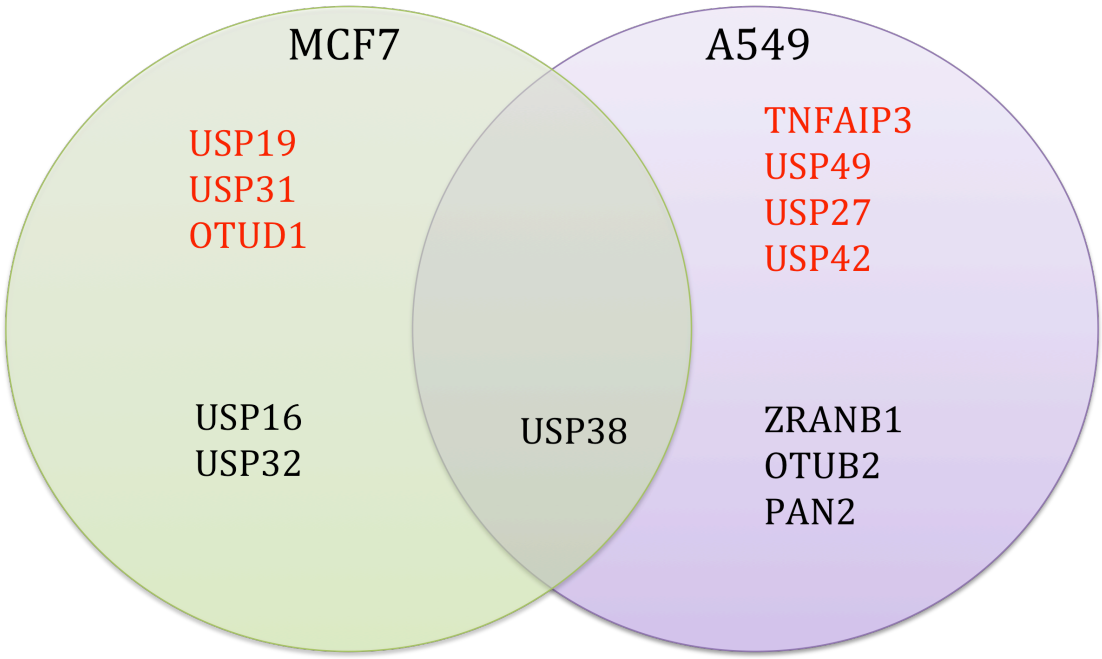

Figure 3.18. Summary of DUBs whose knockdown result in change in E-cadherin level. Summary was based on results of siRNA DUB library screen performed in MCF7 and A549 (presented in Figure 3.13 and Figure 3.17 respectively). Only DUBs, which are identified as top hits by both quantification results are listed. DUBs whose knockdown resulted in increase in E-cadherin level are highlighted in red while those resulted in decrease in E-cadherin level are highlighted in black.

Figure 3.18 summarises the top hits, based on their effect on E-cadherin stability, from the two separate siRNA DUB library screens in MCF7 and A549 respectively. USP38 is the only DUB which when siRNA depleted results in decrease E-cadherin level, while there was no overlapping candidates for DUBs which when siRNA depleted results in increase in E-cadherin level.

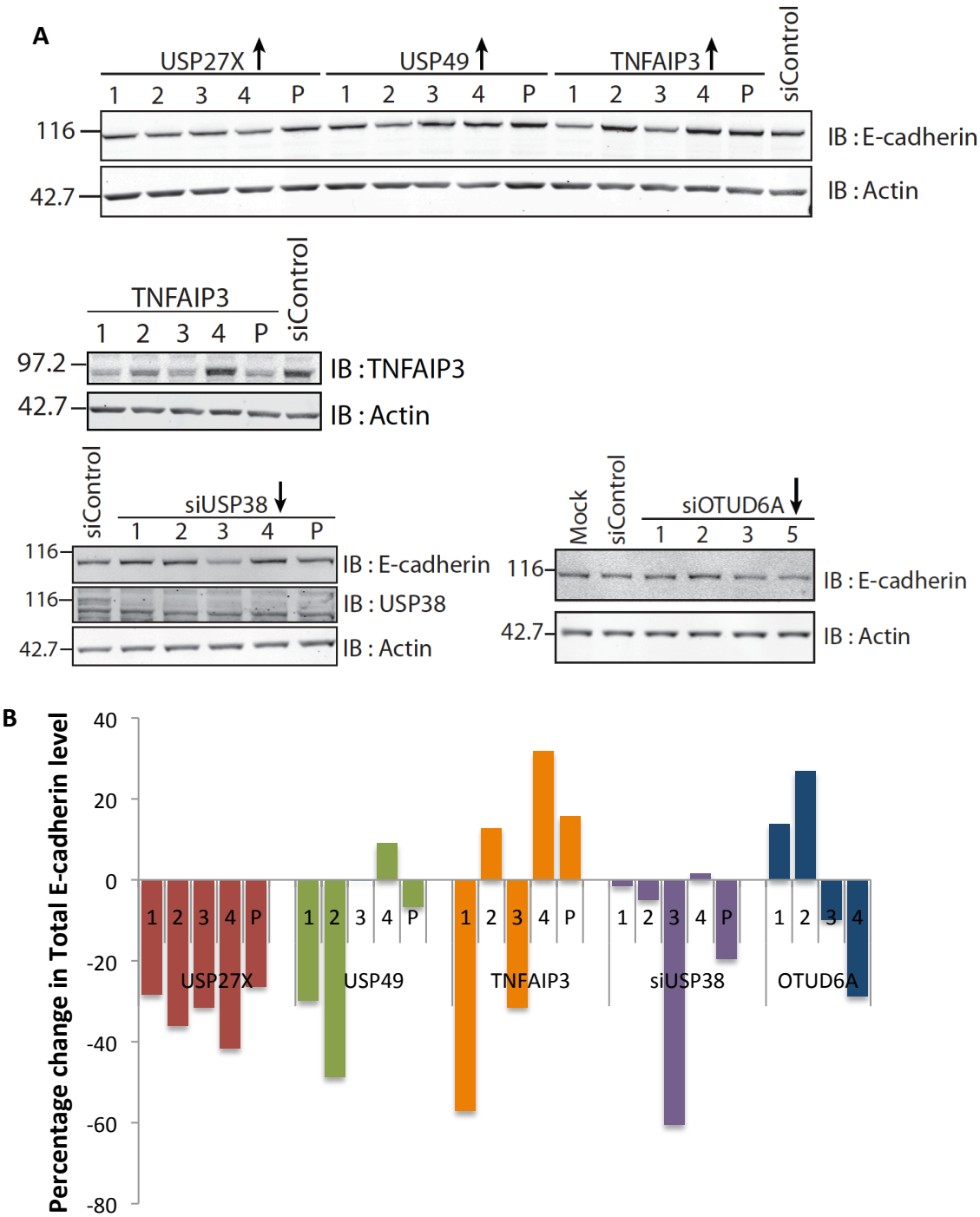

Figure 3.19. Deconvolution of targets which resulted in change in E-cadherin level in A549. (A) A549 cells were transfected with single oligos or pool oligos against candidate DUBs at a final concentration of 40nM. Cells were lysed using NP40 lysis buffer, and lysates were resolved by SDS-PAGE and immunoblotted for E-cadherin. Gel images were acquired by an infra-red scanner (Odyssey, LICOR). Arrows indicate direction of change of E-cadherin following knockdown of corresponding DUBs in the screen. **(B)** Densitometric analysis was performed using Image J software to determine E-cadherin band intensity and normalised to actin level. Change in total E-cadherin level from that of non-targeting controls was then determined and represented in the graph.

Among the targets identified in this screen, TNFAIP3, USP27X, USP49, USP38 and OTUD6A were chosen for deconvolution (Figure 3.19). While the first three candidates were chosen for the dramatic increase in E-cadherin following their depletion, the latter two were chosen because USP38 depletion resulted in the decrease in total E-cadherin level while OTUD6A depletion decreased the full length E-cadherin to 80kDa fragment ratio in MCF7 cells. For both USP27X and USP49, the repeat knockdown using pool oligos did not reproduce the dramatic increase in E-cadherin level that was observed during the screen. All the individual oligos against USP27X and 3 out of 4 oligos against USP49 did not result in any increase, but decrease, in E-cadherin level either. For TNFAIP3, oligo 2 and 4 recapitulated the pool knockdown effect, resulting in increase in E-cadherin level, but not to the extent that was observed with the screen. While 2 oligos giving the same effect as the pool oligos is sufficient to eliminate the possibility of off-target effects, the immunoblot to assess knockdown efficiency of TNFAIP3 suggested otherwise. Firstly, oligo 4 did not effectively silence TNFAIP3 and interestingly, the other 2 oligos, which had higher silencing potency, in fact resulted in decrease in E-cadherin level. For USP38, the decrease in E-cadherin level following transfection of the pool oligos was quite mild, and only oligo 3 resulted in a more dramatic loss of E-cadherin. 2 out of the 4 oligos against OTUD6A resulted in a slight decrease in E-cadherin level. Therefore, the deconvolution results for USP38 and OTUD6A in USP38 were not convincing enough to indicate a functional role between the DUBs and E-cadherin in A549 cells.

3.2.5 Understanding the Functional Relationship between USP38 and E-cadherin

Among the DUBs that were chosen for deconvolution, USP38 was followed up as its depletion resulted in a dramatic change in the E-cadherin ratio and a significant reduction in the total E-cadherin level in MCF7 cells.

A QPCR experiment was performed to measure the level of E-cadherin mRNA. The loss of E-cadherin at the protein level was accompanied by an increase in

E-cadherin mRNA level (Figure 3.20), indicating USP38 does not regulate E-cadherin at transcriptional level. In fact, an inverse correlation was observed between the E-cadherin protein and mRNA level: in the cases where amount of E-cadherin protein was reduced (induced by oligo 1, 2 and 4), an increase of >30% in E-cadherin mRNA level was detected; while when the E-cadherin protein level increased, E-cadherin mRNA level was reduced.

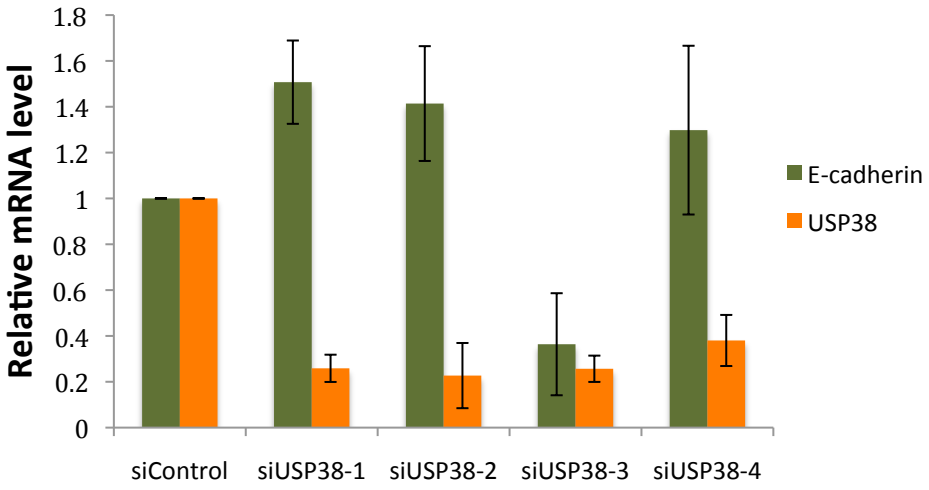

Figure 3.20. USP38 does not regulate E-cadherin at transcriptional level. MCF7 cells were transfected with single oligos against USP38 at a final concentration of 40nM. RNA was extracted 72 hours post transfection and cDNA was prepared by reverse transcription. QPCR was performed using an iQ5 Real-Time PCR system (Biorad), and intron spanning primer pairs were used to measure the mRNA level of E-cadherin, USP38 and actin. All data were normalised to actin level and data shown represents results of 2 biological repeats. Error bars represent standard deviation of data.

The stability of E-cadherin was then checked following siRNA depletion of USP38 by treating MCF7 cells with cycloheximide, which inhibits protein synthesis (Ennis and Lubin, 1964), for different lengths of time. Cycloheximide treatment resulted in decrease in E-cadherin level over time, whether or not USP38 was depleted (Figure 3.21A and B). siRNA depletion of USP38 in MCF7 cells resulted in decrease in E-cadherin level compared to cells transfected with non-targeting siRNA oligo. However, the rates of reduction of E-cadherin level following cycloheximide treatment were the same for both control and USP38-depleted cells (Figure 3.21C). At face value, this indicates that USP38 loss did not affect stability of E-cadherin.

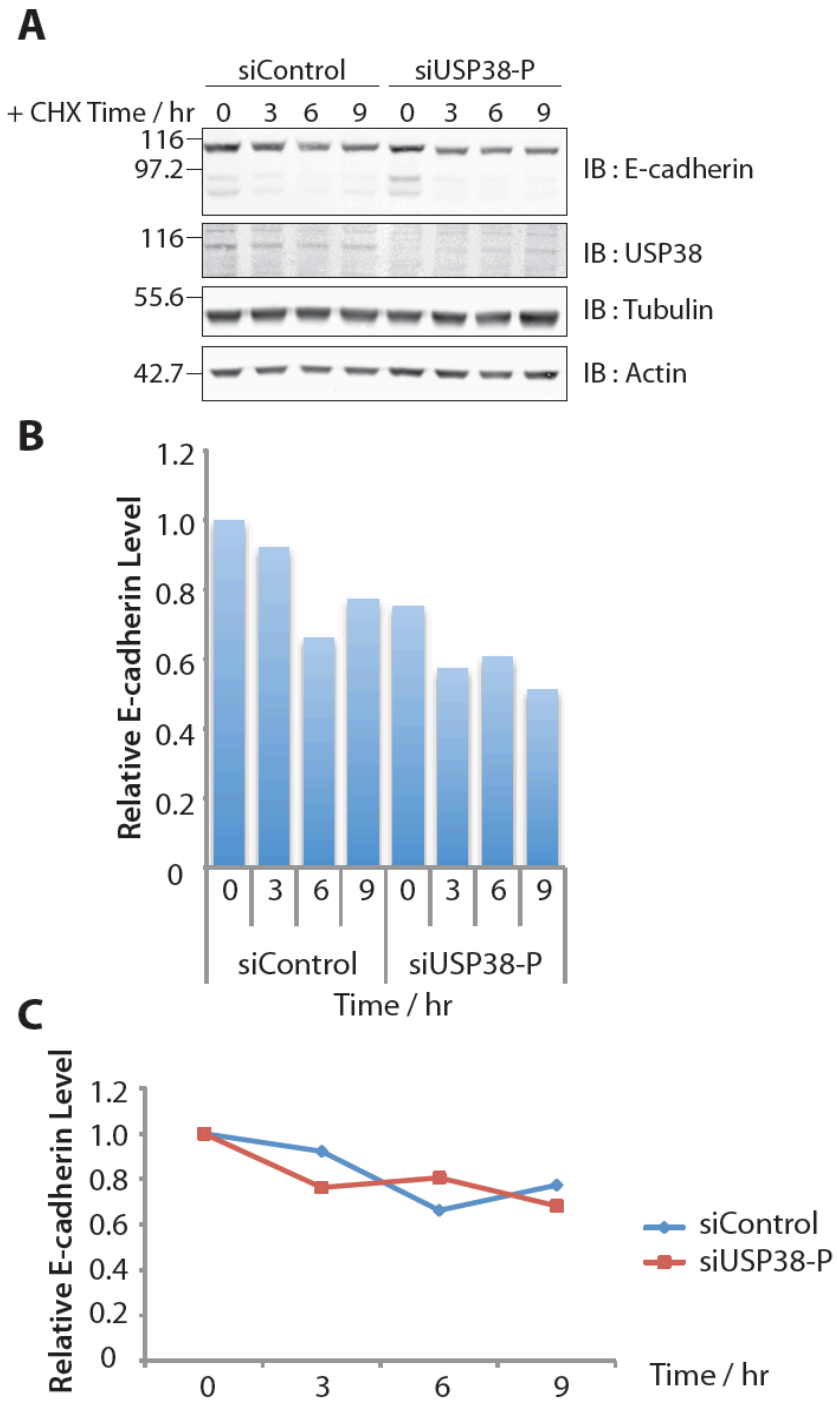

Figure 3.21. Loss of USP38 did not alter rate of E-cadherin degradation. MCF7 cells were transfected with non-targeting siRNA (siControl) or pool of oligos against USP38 at a final concentration of 40nM for 72 hours. Cells were then treated with cycloheximide (CHX) for indicated lengths of time, after when cells were lysed with NP40 lysis buffer. **(A)** Cell lysate was then resolved on 4-12% SDS-PAGE and transferred to nitrocellulose membrane. The membrane was then immunoblotted with antibodies for E-cadherin, USP38, tubulin and actin. **(B)** and **(C)** Densitometric analysis of blot image in (A) was performed to determine relative E-cadherin level, normalised to tubulin, at different time points. Data presented represent average of 2 biological repeats.

As an alternative approach to siRNA knockdown, I transfected MCF7 cells with GFP-tagged USP38 plasmid, to check if E-cadherin level would change in the opposite direction in comparison to USP38 depletion. Overexpression of USP38 did not seem to alter the level of E-cadherin, nor did it cause any significant change in the full length E-cadherin to 80kDa E-cadherin fragment ratio (Figure 3.22).

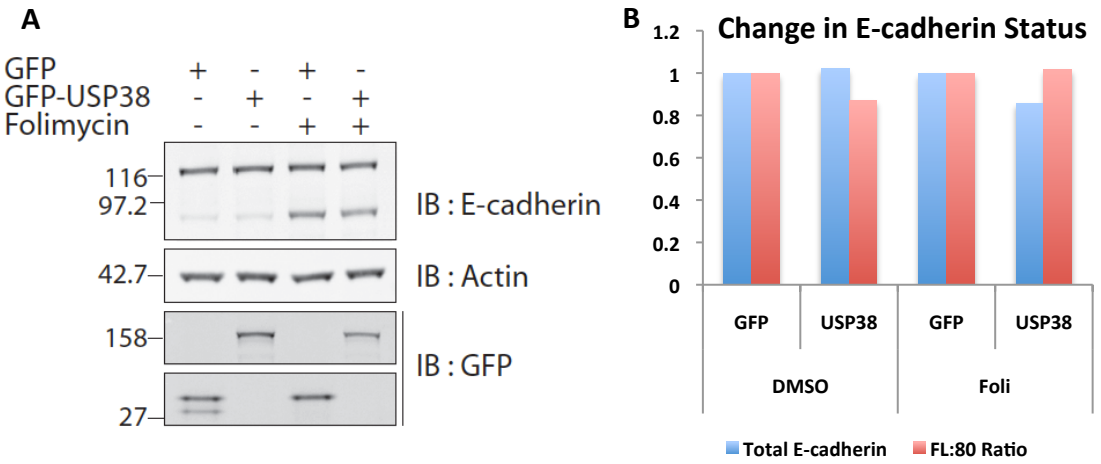

Figure 3.22. Overexpression of USP38 does not affect E-cadherin status. (A) MCF7 cells were transfected with 0.5µg GFP or 1µg GFP-USP38 plasmid, and cells were treated with 10µl DMSO or 100nM Folimycin for 6 hours 24 hours post-transfection. Cells were lysed using NP40 buffer and resolved by SDS-PAGE, followed by immunoblotting. Gel images were acquired by infra-red scanner (Odyssey, LICOR). **(B)** Densitrometric analysis was performed using ImageJ to determine total E-cadherin level and ratio between full length and 80kDa E-cadherin. Transfection efficiency was between 30-35%.

Since E-cadherin is down-regulated following siRNA depletion of USP38, it was expected that the cell-cell adhesion would be affected. I therefore decided to perform a "hanging drop cell aggregation" assay (Figure 3.23) to measure cell-cell contact assembly (Nola *et al.*, 2012). Prior to assessing the effect of USP38 depletion on formation of cell-cell contact, I conducted an experiment to determine the optimal cell number for this assay. At the starting of the experiment, the cells were more dispersed and segregated from each other in small groups. For droplets of higher cell densities, 7.5 x 10^4 cells/ml and 10 x 10^4 cells/ml, the cells were very crowded that most cells were in contact with other cells. For all the different cell

densities, the cells were moving towards the centre of the droplet and aggregated with each other over time.

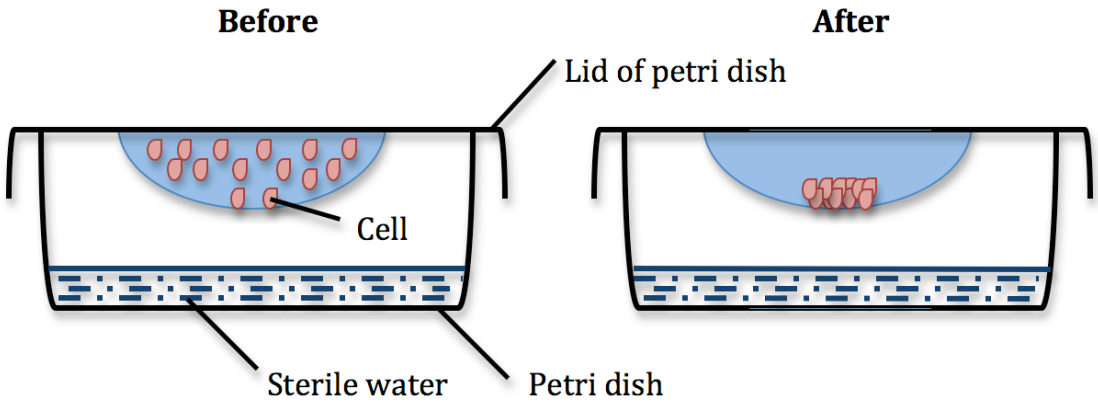

Figure 3.23. Schematic diagram of "hanging drop cell aggregation" assay. Monolayer of adherent cells were disaggregated with 2mM EDTA/PBS and resuspended in full DMEM. 20µl of cell suspension was pipetted onto the inner side of lid of a petri dish and the lid was inverted onto a petri dish containing sterile water ("Before"). The whole set up was incubated at 37˚C for specific length of time so that cells were allowed to fall to the tip of the droplet by gravity and aggregate ("After"). Note: Image was not drawn to scale.

I then decided to seed cells at density of 5 x 10^4 cells/ml since at this cell density, cells were not clumped altogether at the initial time point and the size of cell aggregate at the end of the experiment was of medium size relative to cell aggregates formed with other cell densities. The depletion of E-cadherin abolished drastically the ability of cells to aggregate, as at the end of the 2-hour incubation, the cells were mostly segregated from each other (Figure 3.25A; see Appendix 3.1 for magnified images at 2 hour time point). The siRNA depletion of E-cadherin was accompanied by a dramatic loss of β-catenin (Figure 3.25B). The siRNA depletion of USP38 resulted in a slight decrease of E-cadherin and more than 50% loss of β-catenin, but the cells were still able to aggregate.

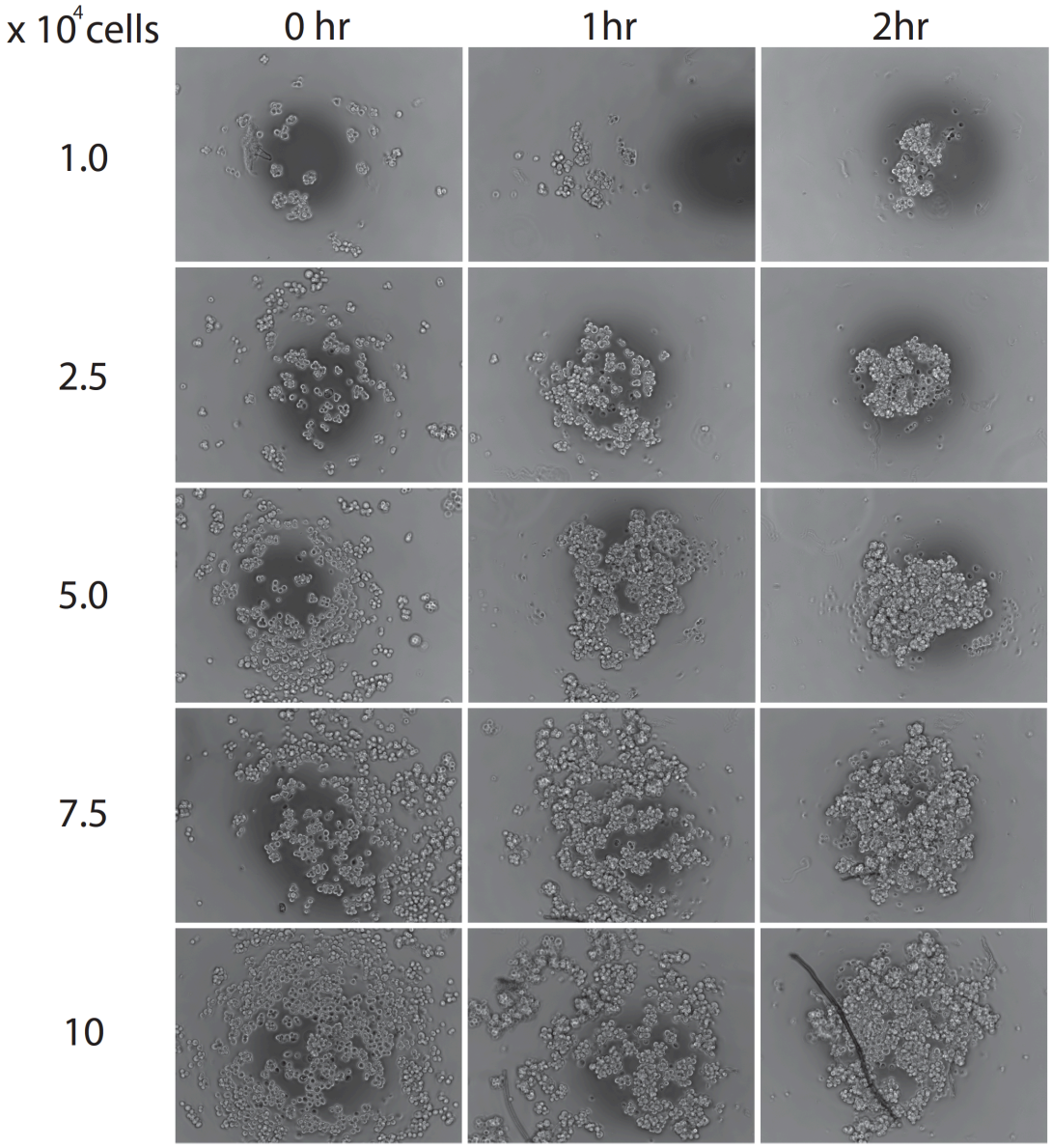

Figure 3.24. Optimisation of cell number for "hanging drop cell aggregation" assay. Monolayer of MCF7 cells grown in 10cm dish were detached by incubating the cells with 5ml of 2mM EDTA/PBS at 37˚C, for 20 minutes. The cells were further disaggregated by pipetting up and down, and cell number was determined using a haematocytometer. After that, cells were resuspended in full DMEM at the following densities: 1 x 10^4 cells/ml, 2.5 x 10^4 cells/ml, 5 x 10^4 cells/ml, 7.5 x 10^4 cells/ml, 10 x 10^4 cells/ml. 20µl of each of the cell suspension was pipetted onto the inner side of lid of a petri dish and the lid was inverted onto a petri dish containing sterile water ("Before"). The whole set up was incubated at 37˚C. Images of cells were acquired at 0 hr, 1 hr and 2 hr.

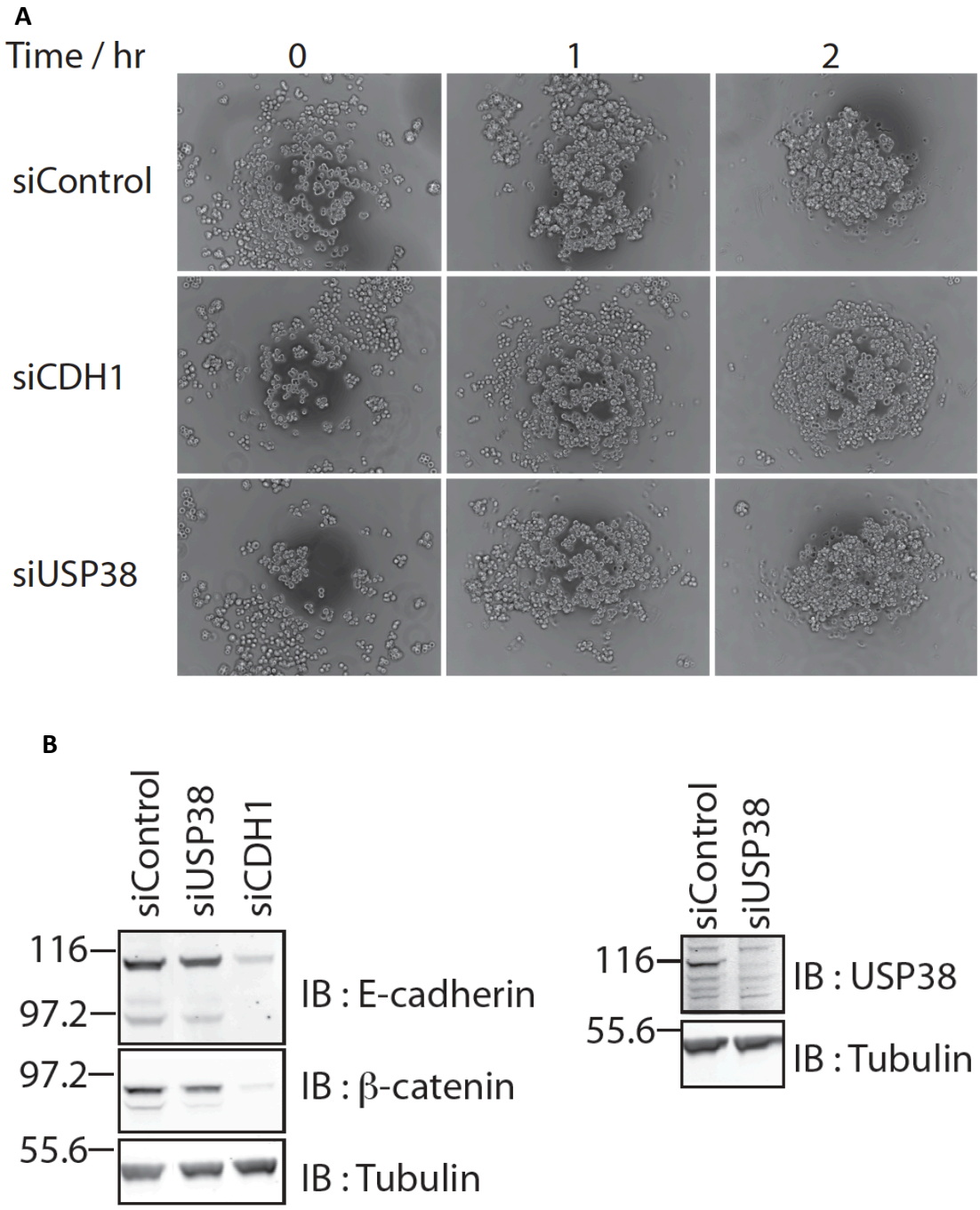

Figure 3.25. USP38 depletion did not abolish ability of cells to aggregate. MCF7 cells were transfected with non-targeting siRNA oligo or siRNA oligo against E-cadherin or USP38, with a final concentration of 40nM for 72 hours. **(A)** After 72 hours, cells were detached from wells by incubation with 2ml of 2mM EDTA/PBS for 20 minutes at 37˚C. The cells were further disaggregated by pipetting up and down, and cell number was determined using a haematocytometer. After that, cells were resuspended in full DMEM at 5 x 10^4 cells/ml. 20µl of the cell suspension was pipetted onto the inner side of lid of a petri dish and the lid was inverted onto a petri dish containing sterile water ("Before"). The whole set up was incubated at 37˚C. Images of cells were acquired at 0 hr, 1 hr and 2 hr. (see Appendix 3.1 for magnified images at 2 hour time point) **(B)** Cells were lysed using NP40 lysis buffer. Lysate was resolved by SDS-PAGE and transferred to nitrocellulose membrane, followed by immunoblotting for E-cadherin, β-catenin, USP38 and tubulin.

3.3 **Discussion**

It has been demonstrated in MDCK (Fujita *et al.*, 2002) and HT29 (Swaminathan and Cartwright 2012) cell lines that HGF stimulation leads to the ubiquitylation of E-cadherin by Hakai, and its subsequent endocytosis. Therefore, at the initial stage of my project, I intended to identify DUB(s), which is/are involved in HGF-dependent regulation of E-cadherin in human cancer cells and began by stimulating a panel of human cancer cell lines with HGF. HGF is known to be able to induce dissociation of epithelial cells (Stella and Comoglio, 1999). Indeed, for all the cancer cell lines (HT29, MCF7, DU145, HCT116) stimulated, the cells were more detached from each other (Figure 3.2). MCF7 and HT29 became more flattened out with relatively round cell shape, and their E-cadherin level remained unchanged after HGF stimulation (Figure 3.1). It was previously reported that HGF stimulation of HT29 and MCF7 led to dissolution of AJ and internalisation of E-cadherin (Reshetnikova *et al.*, 2007; Matteucci *et al.*, 2006). Based on observation of my experiments, although dissolution of AJ was observable following HGF stimulation of these cells, the cytoplasmic distribution of E-cadherin was less distinct. For the case of MCF7, the different observations could be due to the difference in length of time when cells were fixed after HGF stimulation. Matteucci and colleagues fixed MCF7 cells following 1 and 2 hours of HGF stimulation, and they observed prominent E-cadherin staining in cytoplasm at 1 hour, and at the second hour, E-cadherin was observed around the nucleus but not in the cytoplasm (Matteucci *et al.*, 2006). As for my experiment, the MCF7 cells were fixed at 3, 6 and 9 hours after HGF stimulation. Moreover, in this study, they used an antibody against the C-terminus of E-cadherin, while I used an antibody against the N-terminal extracellular domain of E-cadherin. Given the loss of E-cadherin staining from the plasma membrane, the lack of detection of E-cadherin in cytoplasm and/or nucleus was probably due to limitation of the antibody used to detect internalised E-cadherin or varying experimental conditions. Despite E-cadherin internalisation, HGF stimulation did not result in loss of E-cadherin. This has been previously reported in MCF7 cells, but in this study, MCF7 was stimulated with HGF for an hour only (Matteucci *et al.*, 2006) while I stimulated the cells for up to 24 hours. On the other hand, DU145 and HCT116 cells

experienced a drastic change in cell shape to become more elongated (Figure 3.2), reminiscent of cells undergoing epithelial-to-mesenchymal transition (EMT), a cellular process which can be triggered by HGF stimulation (Thiery, 2003; Hube *et al.*, 2005). The E-cadherin level of both of these cell lines were reduced after 24 hours of HGF stimulation (Figure 3.1), in agreement with previous reports (Miura *et al.*, 2001; Saitoh *et al.*, 2009). At least based on the observations of the response of these cell lines to HGF, there was a correlation between E-cadherin level and change in morphology of cells following HGF stimulation: in cells where E-cadherin protein was reduced, the cells underwent more drastic change in cell morphology. Indeed, loss of E-cadherin is considered one of the hallmarks of cells having undergone EMT (Cano *et al.*, 2000).

By treating cells with Folimycin, a lysosomal vacuolar ATPase inhibitor (Muroi *et al.*, 1993), it was revealed that E-cadherin is constitutively degraded via the lysosomal pathway in MCF7, as suggested by the stabilisation of an 80kDa E-cadherin fragment (Figure 3.3). This is also supported by the colocalisation of E-cadherin with two lysosomal markers, namely LAMP1 and CD63, following Folimycin treatment, indicating accumulation of E-cadherin in these compartments due to blockage of degradation. It has been previously reported that in a confluent layer of MDCK cells, E-cadherin is endocytosed and recycled back to the plasma membrane under physiological conditions (Le *et al.*, 1999) and it was also demonstrated in isolated MCF7 cells that unbound E-cadherin on cell surface undergoes constitutive internalisation mainly via a clathrin-independent pathway (Paterson *et al.*, 2003), but constitutive degradation of E-cadherin under steady-state condition has never been reported. Since E-cadherin staining was still retained on plasma membrane following Folimycin treatment (Figure 3.5), this suggests that the recycling of E-cadherin to plasma membrane was not impaired. Moreover, the unchanged level of full length E-cadherin (Figure 3.3) suggests that the continual loss of E-cadherin by lysosomal degradation was replenished by continual synthesis of E-cadherin. It is known that the trafficking of membrane associated protein via the endo-lysosomal pathway is a ubiquitin-dependent process (Clague *et al.*, 2012) and this is also likely the case for E-cadherin, which has also been shown to undergo ubiquitin-dependent

internalisation. However, I did not assess ubiquitylation status of E-cadherin in my study.

I then decided to use the relative level of the full length E-cadherin and 80kDa E-cadherin fragment as biochemical readout for a siRNA DUB library screen. One added benefit of measuring the ratio of full length E-cadherin to the 80kDa E-cadherin fragment is that both of the bands are recognised by the same antibody at the same affinity, and it does not require normalisation to tubulin or actin. I have done 3 technical replicates of the Western blot analysis, which is necessary. Although Western Blotting is useful in detecting changes in protein levels, there are several major sources of errors to this technique, including quality of polyacrylamide gels used (for instance, gel heterogeneity can lead to unequal protein electrophoresis and protein transfer), efficacy and consistency of protein transfer to nitrocellulose membrane, composition of buffers used, procedures of setting up transfer and quality of antibodies used (Aksamitiene *et al.*, 2007). These problems can be augmented by use of large gels for high-throughput studies, which was also the case for my experiments. Therefore, more technical replicates of Western blot analysis are essential to allow statistically significant quantitation and therefore to increase confidence of an observed change. However, a more drastic change in protein level is less sensitive to these technical issues. For example, in the siRNA DUB screen performed in A549, knockdown of USP27X, USP49 and TNFAIP3 led to very significant change in E-cadherin level, which was highly reproducible between technical replicates (Figure 3.16). In order to confidently choose potential targets for deconvolution, I have also used 3 different statistical approaches to analyse the data and selected some candidates for deconvolution.

USP4, USP14, USP41, OTUD6A, STAMBP and STAMBPL1 were the DUBs, which when depleted, resulted in decrease in full length E-cadherin to 80kDa fragment ratio. None of these DUBs has a clear role in regulation of AJ components. USP4 has been implicated in regulation of diverse cellular processes, including splicing of mRNA (Song *et al.*, 2010) and regulation of signaling pathways including NF-κB signaling (Fan *et al.*, 2011; Xiao *et al.*, 2012), TGF-β signaling (Zhang *et al.*,

2012) and Wnt signaling (Zhao *et al.*, 209). USP41 has been reported to be a pro-apoptotic DUB as its overexpression leads to activation of caspase-3 and downregulation of p21, p27 and cyclin B, which are positive regulators of the cell cycle (Gewies and Grimm, 2003). The biological function of OTUD6A is unknown.

USP14 is a proteasome-associated DUB and is important for the disassembly of polyubiquitin chain (Borodovsky *et al.*, 2001). Using an in vitro deubiquitylation assay, Hu *et al.* demonstrated that USP14 removes ubiquitin from the distal end of a K48-linked polyubiquitin chain (Hu *et al.*, 2005). There is growing evidence suggesting that K63-linked ubiquitin chain is the main chain type for sorting into multivescular body and lysosome (Duncan *et al.*, 2006; Lauwers *et al.*, 2009). If that is also true for the case of E-cadherin lysosomal degradation, the change in the full length to 80kDa E-cadherin fragment ratio may not be a direct effect of loss of USP14.

STAMBP, signal transducing adaptor molecule (Swanson, Kang et al.) binding protein, also known as AMSH, is a DUB localises to the endosomal compartments (McCullough *et al.*, 2004) and shows linkage specificity for K63-linked polyubiquitin chains (McCullough *et al.*, 2006). siRNA depletion of STAMBP accelerates degradation of epidermal growth factor receptor (Siegfried, Chou et al.) following EGF stimulation, suggesting its negative regulatory role in ubiquitin-dependent sorting of EGFR to lysosomes (McCullough *et al.*, 2004). The overexpression of catalytically inactive mutant of STAMBP results in accumulation of ubiquitylated proteins on endosomes, suggesting a more general role of STAMBP in regulating sorting of ubiquitylated cargo. If this regulatory role of STAMBP on EGFR is extendable to other ubiquitylated cargo, such as E-cadherin, it explains the observed decrease in full length E-cadherin to 80kDa E-cadherin fragment ratio, as an increased ratio indicates that E-cadherin degradation is favoured. STAMBPL1, also known as AMSH-LP, is a close homolog of STAMBP. Similar to STAMBP, it also localises to the endosomal compartment, owing to a clathrin-binding sites conserved between the two proteins, and overexpression of catalytically inactive mutant of STAMBPL1 resulted in accumulation of ubiquitylated proteins on endosomes

(Nakamura *et al.*, 2006). This evidence suggests functional redundancy of STAMBP and STAMBPL1 in regulation of ubiquitylated cargos on the endosome. (However, STAMBP and STAMBPL1 are not functionally redundant in regulation of TGFβ signaling based on unpublished research data by Monika Chojnowska, an ex-colleague in our laboratory). Therefore, it is not unexpected that their siRNA depletion resulted in similar change, i.e. decrease, in the full length E-cadherin to 80kDa E-cadherin fragment ratio.

USP8, USP12, USP38 and USP40 were the DUBs, which when siRNA depleted, resulted in increase in full length E-cadherin to 80kDa E-cadherin fragment ratio (Figure 3.12). Among these, effect of USP12 and USP40 knockdown on E-cadherin ratio was confirmed to be off-target effect. USP8 is also an endosomal associated DUB and it has been demonstrated in our lab that siRNA depletion of USP8 resulted in delayed degradation of EGFR upon EGF stimulation (Row *et al.*, 2006). This may happen as a result of a sorting defect due to loss of STAM, which is stabilised by USP8. Such a regulatory mechanism may account for the increase in full length E-cadherin to 80kDa E-cadherin fragment ratio.

Among the targets that were deconvoluted, I decided to focus on USP38. This was because the siRNA depletion of USP38 resulted in the most dramatic increase in the full length E-cadherin to 80kDa E-cadherin fragment ratio, accompanied by a significant loss of E-cadherin. Moreover, USP38 was an unstudied DUB that I wanted to characterise (See Chapter 6). The loss of E-cadherin protein following knockdown of USP38 was not accompanied by a decrease in E-cadherin mRNA (Figure 3.20). In fact, the E-cadherin mRNA level increased following USP38 knockdown. This suggests USP38 does not exert its regulatory role on E-cadherin at the transcriptional level. The loss of E-cadherin protein possibly activates a feedback mechanism, which attempts to counteract the loss by promoting the gene expression of E-cadherin. While DUBs can enhance stability of its substrate by rescuing the protein from proteasomal or lysosomal degradation, this was also not the case for USP38 and E-cadherin, since siRNA depletion of USP38 did not lead to enhanced E-cadherin turnover (Figure 3.21). In other words, USP38 does not regulate E-cadherin at post-

translational level. However, it is worth noting that this experiment used Cycloheximide to inhibit protein translation and the inhibitor may lead to undesirable effects to other processes within cells, such as protein degradation. Nevertheless, putting the results of the two experiments together (Figure 3.20 and Figure 3.21), there is only one possibility left for the effect of USP38 knockdown on E-cadherin, which is the loss of USP38 negatively impacts on E-cadherin translation. This is a likely explanation since USP38 was found interacting with two ribosomal proteins, namely ribosomal protein L7 (RPL7) and 40S ribosomal protein S12 (RPS12) (Sowa *et al.*, 2009), implicating a role of USP38 in regulation of the translational machinery. In that case, the effect of USP38 knockdown may not be specific to E-cadherin, but represents a general defect in protein translation. I have however not tested this hypothesis during the course of my research project.

I was not able to verify further the functional relationship between E-cadherin and USP38. If USP38 is acting directly on E-cadherin and rescues it from lysosomal degradation, it would be expected that following its overexpression, both E-cadherin level and the ratio of the full length to 80kDa E-cadherin fragment to be increased. The fact that both of these measurements were not altered (Figure 3.22) therefore suggests that the concentration of USP38 is not limiting in regulation of E-cadherin. Another possible explanation is that USP38 does not exert its regulatory role directly on E-cadherin This may be supported by the observation that β-catenin was more significantly reduced than E-cadherin following USP38 knockdown (Figure 3.25B). Since both E-cadherin and β-catenin are co-regulated in the AJ complex formation (Chen *et al.*, 1999), it is possible that the reduction in E-cadherin level was secondary to the loss of β-catenin.

While loss of E-cadherin following USP38 knockdown was apparent, such extent of loss was not sufficient to abolish AJ formation and cell aggregation. This suggests expression level of E-cadherin on the plasma membrane was maintained despite the loss of E-cadherin, probably by enhancing E-cadherin recycling as indicated by the increased full length to 80kDa E-cadherin ratio. Only when E-cadherin was significantly depleted, the ability of cells to aggregate was abrogated.

3.4 Conclusion

In conclusion, through the works presented in this chapter, I demonstrated that E-cadherin undergoes constitutive turnover via the lysosomal pathway in MCF7 cells. Such mode of E-cadherin regulation was however not seen in other cell lines tested. In the effort to identify DUBs regulating E-cadherin, I found USP38, which affects indirectly turnover of E-cadherin protein. The mechanism of how these two proteins are functionally related was however not uncovered. To understand this, one possible direction of future work is to assess effect of USP38 knockdown on other proteins known to play a role in regulation of AJ formation.

Chapter 4

Identification of Deubiquitylases Involved in Regulation of β–catenin

4.1 Introduction

β-catenin, together with α- and γ-catenin, was first identified by Ozawa and colleagues in 1989 as a cytoplasmic binding partner of E-cadherin, providing a physical anchorage of the E-cadherin molecules on the plasma membrane to the actin cytoskeleton (Ozawa *et al.*, 1989; Ozawa *et al.*, 1990). A few years later, two independent groups discovered the interaction of β-catenin with the tumor suppressor gene, adenomatous polyposis coli (APC) (Rubinfeld *et al.*, 1993; Su *et al.*,1993) and Nusse's group demonstrated its role in mediating Wnt signaling (Noordermeer *et al.*, 1994). This dual role of β-catenin in two different cellular processes, cell-cell adhesion and development, which are both related to cancer development, has drawn immense research interest to understand its regulation.

It is now know that the tight regulation of cytoplasmic level of β-catenin protein is under the control of the ubiquitin-proteasome system (UPS) (Aberle *et al.*, 1997; reviewed in Logan and Nusse, 2004). In fact, there is burgeoning evidence showing the role of UPS as a master regulator of Wnt signaling at multiple levels,

through its regulation of different protein components of the pathway. For example, K48-linked ubiquitylation of APC (Choi *et al.*, 2004), Dishevelled (Angers *et al.*, 2006) and Axin (Lui *et al.*, 2011) targets them for proteasomal degradation whereas K63-linked ubiquitylation of APC (Tran *et al.*, 2008) and Dishevelled (Tauriello *et al.*, 2010) affects their activity in transducing Wnt signaling cascades. The DUBs, which can deubiquitylate these proteins have also been identified, namely Trabid for APC (Tran *et al.*, 2008), CYLD for Dishevelled (Tauriello *et al.*, 2010) and USP34 for Axin (Lui *et al.*, 2011). While the regulation of these Wnt signaling components by the UPS can affect β-catenin regulation indirectly, the DUB which directly regulates β-catenin is yet to be identified.

The work presented in this chapter aims (i) to identify DUB(s) involved in β-catenin regulation and (ii) to characterise the functional relationship between the DUB(s) and β-catenin.

4.2 Results

4.2.1 Identification of BAP1 as a potential DUB regulating β-catenin

To identify DUBs involved in the regulation of β-catenin at steady state, I have used the large-scale DUB library screen A549 lysate (described in section 3.2.4) for SDS-PAGE and Western Blot analysis. 2 96-well plates of lysates were run and Figure 4.1 shows a representative blot image acquired for this set of experiments. Across the gels, there were multiple samples with higher levels of β-catenin relative to control, such as the knockdown samples of UCHL3, USP6, USP27X, USP54 and TNFAIP3. On the other hand, significant loss of β-catenin following DUBs knockdown was less obvious as judged by eye.

Densitometric analysis of the blot images was performed to give a quantitative measurement of the β-catenin level (Figure 4.2). The DUBs, whose

knockdown resulted in at least a 2-fold increase in β-catenin level were USP27X, TNFAIP3, USP54 and USP32. Among these, the effect of USP27X and TNFAIP3 depletion was the most dramatic, resulting in about a 4-fold increase in β-catenin level. On the other hand, siRNA depletion of BAP1, USP9Y, DUB4, PSMD7, STAMBPL1 and MPND resulted in at least a 2-fold decrease in β-catenin level.

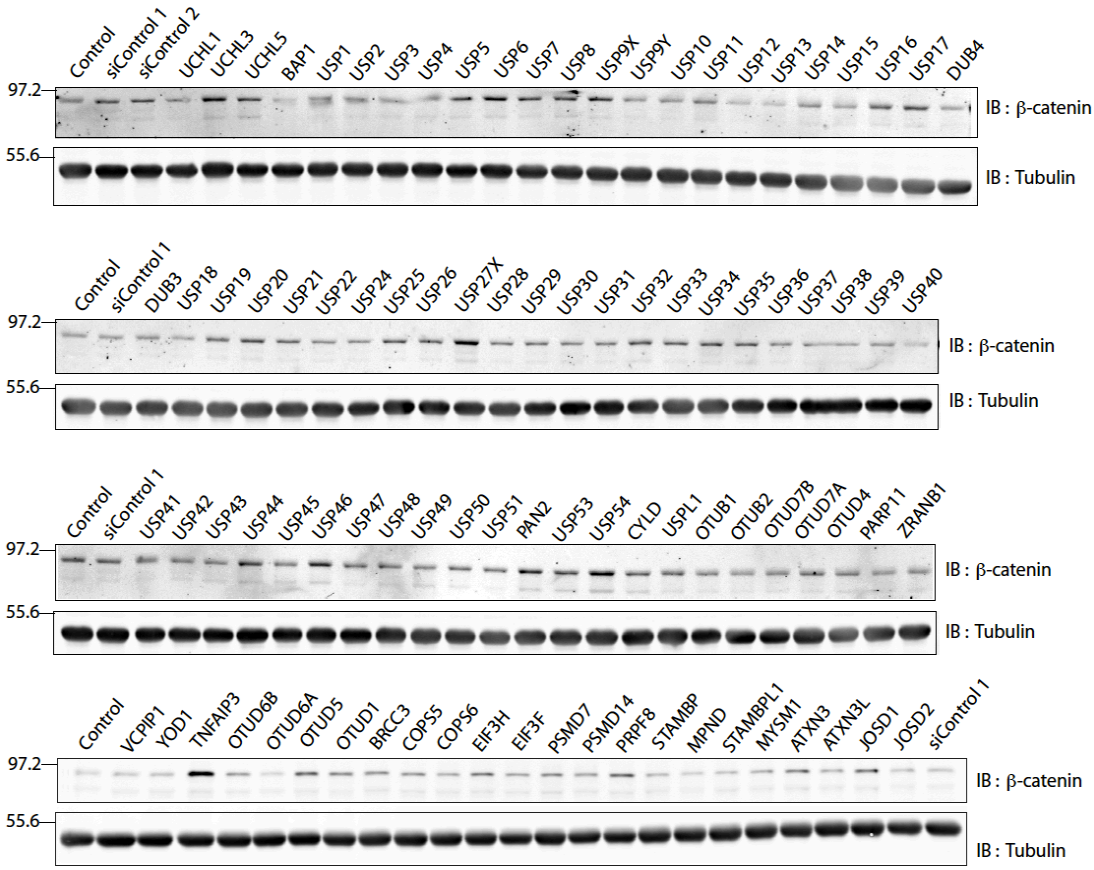

Figure 4.1. **siRNA DUB library screen to identify DUBs regulating β–catenin in A549 cells.** NP40 samples from a large scale A549 DUB screen (Figure 3.15) were resolved on 10% SDS-PAGE gels and transferred to nitrocellulose membrane before immunoblotting with antibody against β-catenin and tubulin. Blot images were acquired by infra-red scanner (Odyssey, LICOR).

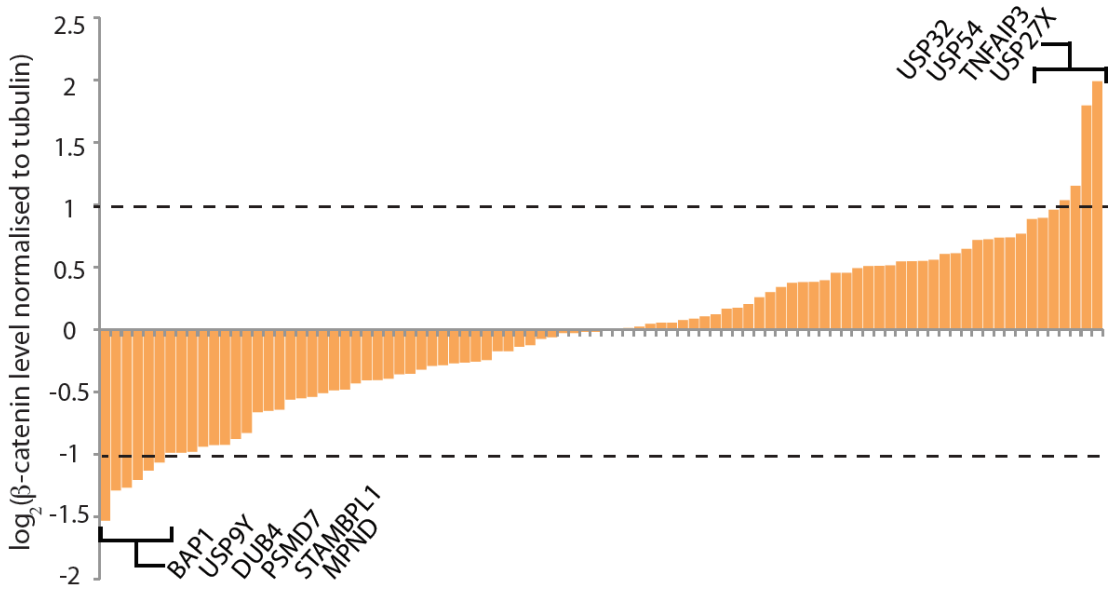

Figure 4.2. Change in β–catenin level following knockdown of DUBs. Densitometric analysis of blot images (Figure 4.1) was performed using ImageJ to determine relative amount of β–catenin level, normalised to the level of tubulin and median. The normalised β–catenin level was then log-transformed and ranked in ascending order. Columns exceeding the dotted line represent samples with more than two-fold change in β–catenin protein level.

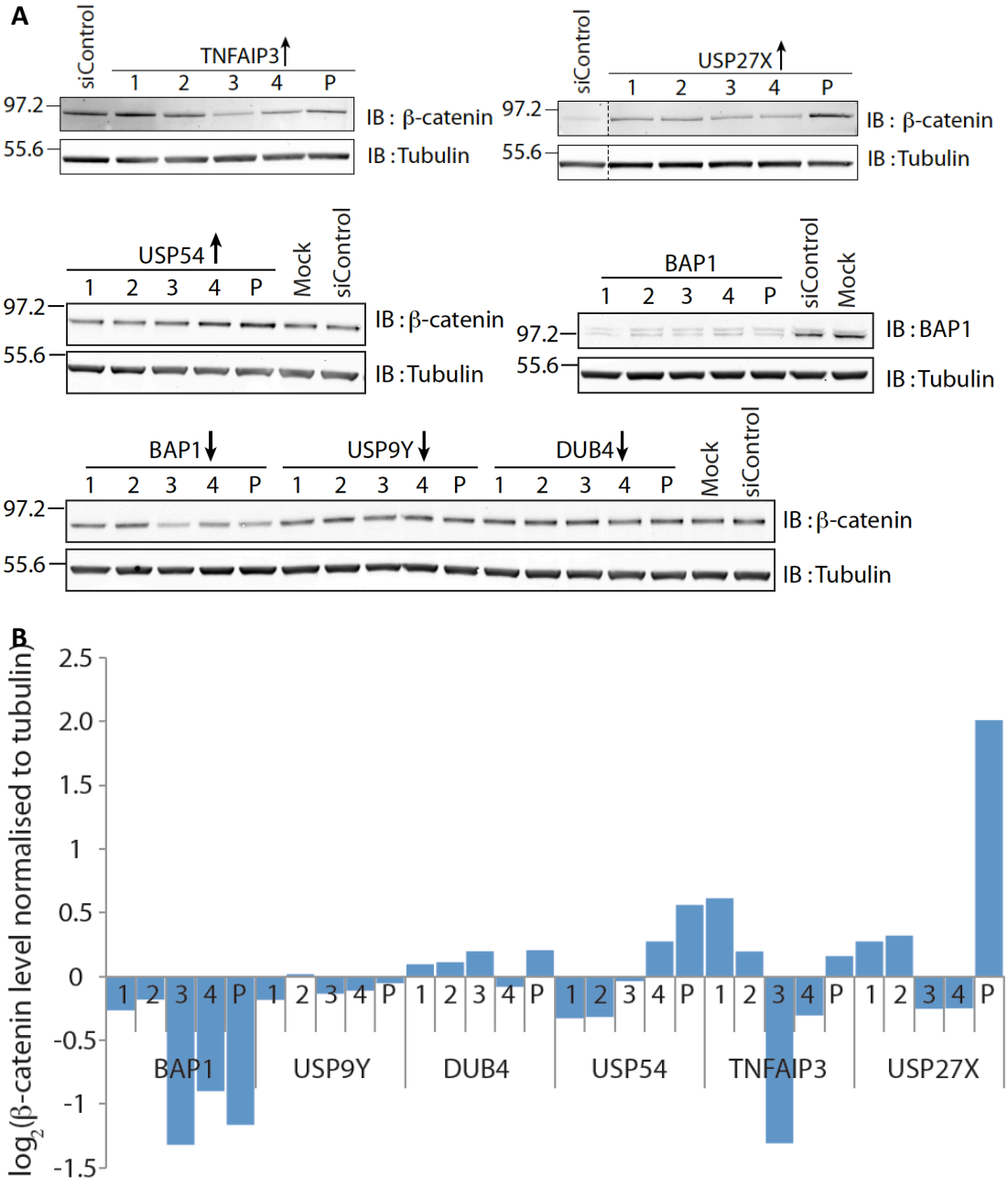

Figure 4.3. **Deconvolution of targets which altered β–catenin level. (A)** A549 cells were transfected with single oligos against candidate DUBs at a final concentration of 40nM. Cell were lysed using NP40 lysis buffer 72 hours later. Lysates were resolved by SDS-PAGE and immunoblotted for β–catenin and tubulin. Gel images were acquired by infra-red scanner (Odyssey, LICOR). Note: The siControl and knockdown samples of USP27X were run on the same gel, and part of the gel was cropped as indicated by the dotted lines. **(B)** Densitometric analysis was performed using Image J to determine the level of β–catenin following normalisation to tubulin level. The normalised β–catenin of each sample was then again normalised to that of the siControl sample and was log transformed. Data shown represents average of 3 technical repeats, and the 5 columns each represents oligo 1, 2, 3, 4 and pool respectively (See Figure 3.18 for knockdown efficiency of TNFAIP3).

From both ends, only the top 3 candidates were chosen for subsequent deconvolution experiments (Figure 4.3). For BAP1, the knockdown using pool oligos reproduced the decrease in β-catenin level as was observed in the screen, and oligos 3 and 4 recapitulated the pool knockdown effect. Immunoblotting with BAP1 antibody confirmed that BAP1 was significantly depleted in all knockdown samples. Among the knockdown performed using single oligos, oligo 3 gave the most dramatic depletion in β-catenin level. Both the USP9Y and DUB4 knockdown using pool oligos did not reproduce the decreased β-catenin level observed in the screen and the β-catenin level was similar to that of the non-targeting siRNA (siControl) sample. The knockdown performed using single oligos against these 2 DUBs also did not result in significant change in β-catenin level compared to the siControl sample.

The knockdown of USP54 using pool oligos reproduced the increased β-catenin level that was observed in the screen. However, this effect of USP54 depletion on β-catenin level was only recapitulated by oligo 4, but not the other 3 oligos, suggesting the increase in β-catenin level was an off-target effect. The knockdown using pool oligos against TNFAIP3 did not reproduce the increased β-catenin level observed in the screen. Among the single oligos against TNFAIP3, only oligo 1 resulted in a slight increase in β-catenin level relative to the siControl (Figure 4.3A) while knockdown using oligos 3 and 4 resulted in depletion of β-catenin. Since all 4 oligos resulted in depletion of TNFAIP3 (Figure 3.18) and that the knockdown effect of TNFAIP3 on β-catenin level was not consistent, the observed increase in β-catenin level following TNFAIP3 knockdown in the screen was very likely an off-target effect. The knockdown of USP27X using pool oligos again resulted in about a 4-fold increase in β-catenin level. However, none of the 4 single oligos recapitulated such an extent of increase: where oligos 1 and 2 resulted in a marginal increase in β-catenin level while oligo 3 and 4 resulted in a marginal decrease in β-catenin level. Knockdown efficiency of USP27X was not checked since the antibody for USP27X was not available in our laboratory. Hence, among the targets that were deconvoluted, only BAP1 emerged as a potential regulator of β-catenin.

4.2.2 Characterisation of functional relationship between BAP1 and β–catenin in MCF7

To check if the regulatory role of BAP1 on β-catenin is unique to A549 cells or not, I then explored the consequence of BAP1 depletion in two other human cancer cell lines, namely MCF7 and SW480 cells. MCF7 cell was chosen because I have performed E-cadherin related experiments in this cell line (see Chapter 3) and given the known function of E-cadherin and β-catenin as core components of the AJ, using the same cell line would allow me to determine the interrelationship between the DUBs idenfied and the two proteins. SW480 cell was chosen because this cell line harbours a truncation mutation in APC, so β-catenin is not efficiently degraded and aberrantly stabilised in this cell line. It would be of therapeutic interest to check if loss of BAP1 function can affect β-catenin in this cell line. The siRNA depletion of BAP1 by all 4 oligos also resulted in decrease in β-catenin level in MCF7 cells (Figure 4.4A), indicating that the regulatory role of BAP1 on β-catenin was not restricted to A549 cells only. However, for the case of BAP1 depletion in SW480 cells (Figure 4.4B), only oligo 1 resulted in a significant depletion of β-catenin, while the other oligos did not result in change in β-catenin level. Therefore, I did not pursue any further with SW480 cells.

When observed under microscope, for A549, the control cells (mock transfected or transfected with non-targeting siRNA), showed prominent β-catenin staining on the plasma membrane at cell-to-cell junction and a strong nuclear staining for BAP1 (Figure 4.5A). For cells transfected with the siRNA oligos against BAP1, the nuclear staining of BAP1 was much weaker, indicating the depletion of BAP1. In these cells, the β-catenin staining on the plasma membrane was much weaker, which is in agreement with the biochemical data (Figure 4.3A).

MCF7 cells, which were mock transfected or transfected with non-targeting siRNA, were tightly bound to each other with prominent β-catenin at cell-to-cell junction and strong nuclear staining for BAP1. Similar to A549 cells, the MCF7 cells transfected with siRNA oligos against BAP1 showed less β-catenin on the plasma

membrane and weaker nuclear staining of BAP1, indicating BAP1 depletion. Moreover, as seen in Figure 4.5B, cells transfected with oligos 2, 3 and 4 were clearly less tightly bound to each other and more flattened out.

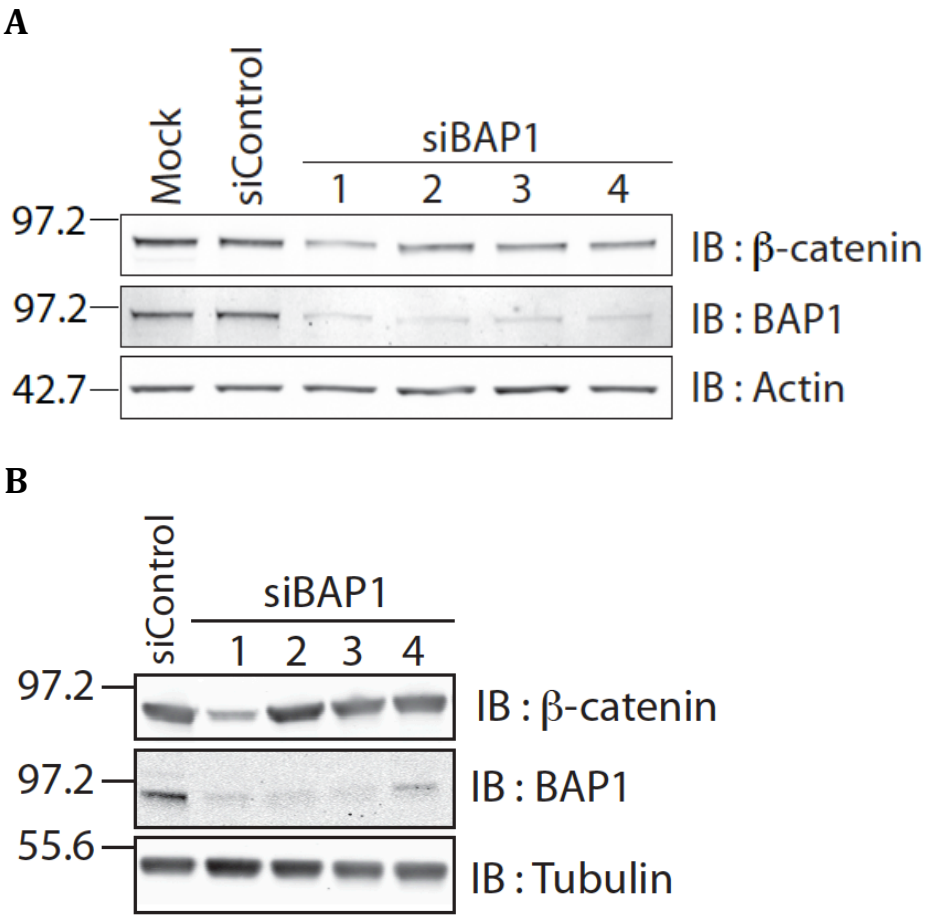

Figure 4.4. BAP1 depletion in MCF7 and SW480 cells. (A) MCF7 and **(B)** SW480 cells were transfected using transfection reagent alone (mock) or non-targeting siRNA oligo (siControl) or siRNA single oligos against BAP1 (40nM) for 72 hours. Cell were lysed using NP40 lysis buffer 72 hours later. Lysates were resolved on 4-12% NuPAGE gel and immunoblotted for β–catenin, BAP1 and actin. Gel images were acquired by infra-red scanner (Odyssey, LICOR).

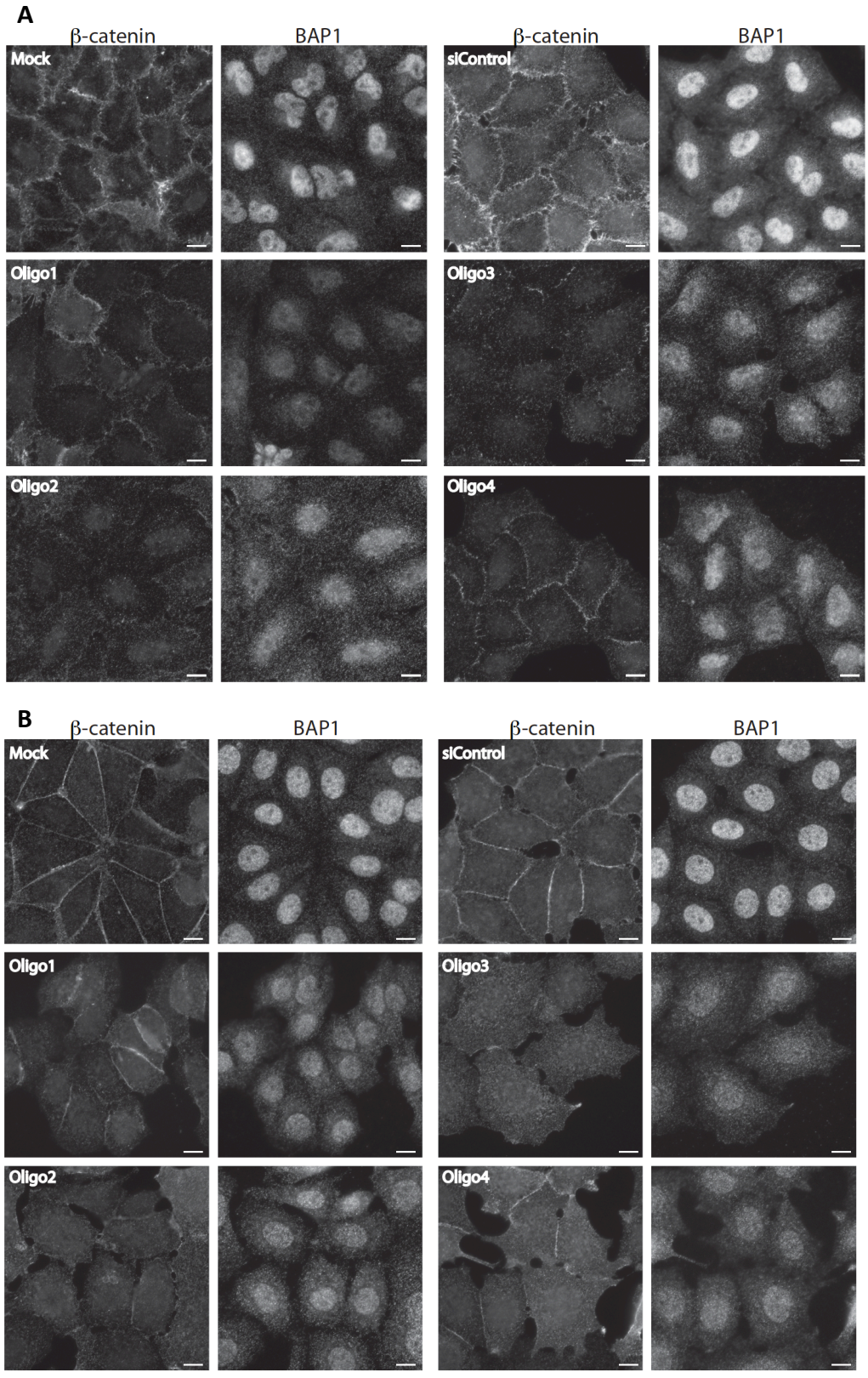

Figure 4.5. BAP1 depletion decreases β–catenin level on plasma membrane in (A) A549 and (B) MCF7 cells. Cells were transfected using siRNA single oligos against BAP1 at 40nM for 72 hours. After that, cells were fixed in 0.4% PFA/PBS, permeabilised and immunostained with antibody against b–catenin and BAP1. Images were acquired using Nikon microscope and the image setting was the same for all images (Scale bar = 10µm).

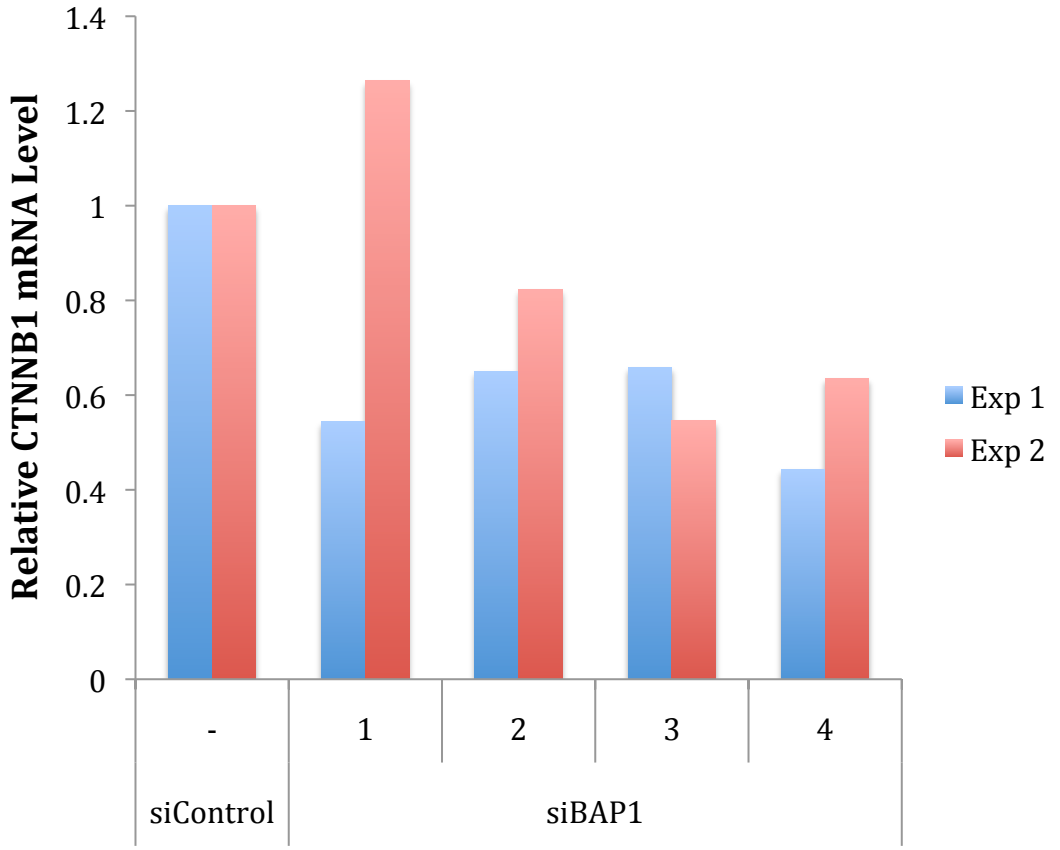

Figure 4.6. siRNA depletion of BAP1 resulted in decrease in β–catenin mRNA level. Cells were transfected using siRNA single oligos (40nM) for 72 hours and mRNA was purified using RNeasy Kit. 1µg of mRNA was reverse-transcribed and subjected to QPCR analysis. Two biological replicates were analysed and QPCR reactions were set up in triplicate for each biological replicate.

I have also measured the mRNA level of β-catenin following BAP1 knockdown in MCF7 and the experiment was done twice. siRNA depletion of BAP1 using oligos 2, 3 and 4 resulted in a decrease of β-catenin mRNA by 20-40%, which was consistent for both biological repeats (Figure 4.6). For knockdown using siRNA oligo 1, at least in one experiment, there was a decrease of β-catenin mRNA by about 50%.

When BAP1 was transiently overexpressed in MCF7 cells, there was a reproducible (the experiment was repeated three times), higher level of β-catenin compared to cells transiently overexpressing GFP (Figure 4.7A). This suggested a positive regulatory role of BAP1 on β-catenin in MCF7. To assess if this is dependent on the catalytic activity of BAP1, I repeated the experiment by including two mutant

of BAP1, namely BAP1-C91S and BAP1-A95D. Both of these mutants lack deubiquitylating activity based on *in vitro* studies (Ventii *et al.*, 2008). The overexpression of these two mutants did not lead to change in β-catenin level, while the overexpression of wildtype BAP1 resulted in a higher level of β-catenin (Figure 4.7B). At face value, it seemed regulatory of BAP1 on β-catenin is dependent on its catalytic activity. However, the expression level of the mutant BAP1 was not as high as that of the wildtype BAP1 and this difference in expression level may account for the lack of effect. In this experiment, I also observed an extra lower molecular weight band below the β-catenin band, which was likely a non-specific band. This is because in the two experiments shown in Figure 4.7, the same antibody was used and β-catenin always runs slightly below the 97.2kDa marker. Moreover, that was the band which showed an increase when wildtype BAP1 was overexpressed, which is a very reproducible effect of BAP1 overexpression.

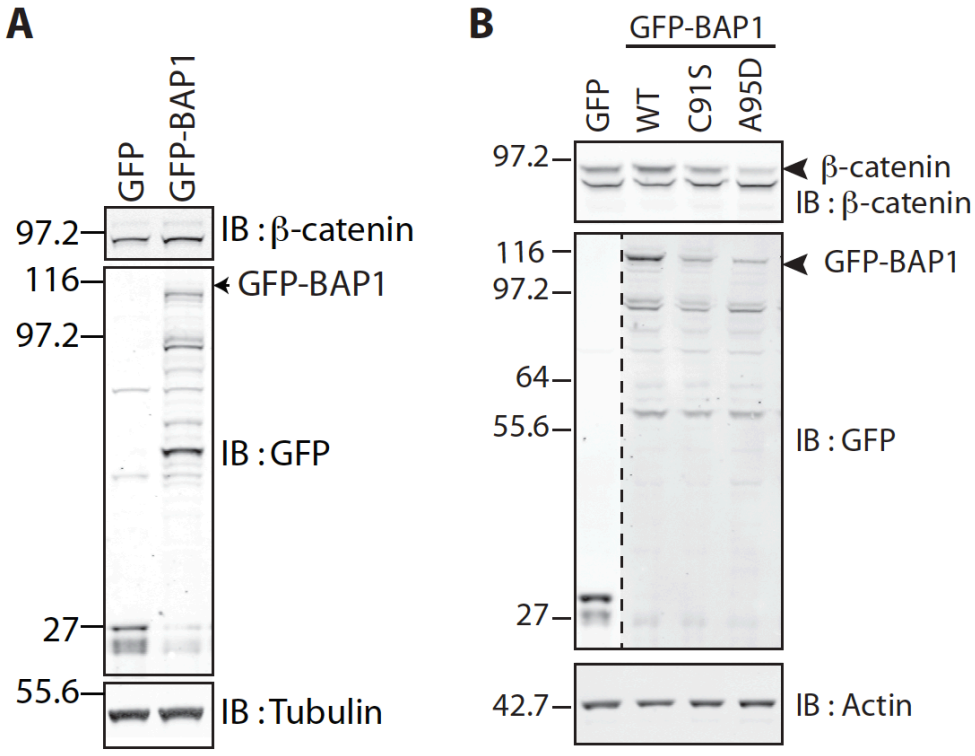

Figure 4.7. Upregulaton of β–catenin following BAP1 overexpression. MCF7 cells were transfected with 0.5μg of GFP empty vector or 1μg of GFP-tagged BAP1, BAP1-C91S and BAP1-A95D fusion plasmids. Cell were lysed using NP40 lysis buffer 24 hours later and lysates were resolved on 4-12% NuPAGE gels and immunoblottded for β–catenin, GFP and actin. Blot images were acquired by infra-red scanner (Odyssey, LICOR). Transfection efficiency was not determined in this experiment.

To further investigate the role of BAP1 on β-catenin regulation, I decided to do a rescue experiment. Before carrying out that experiments, I had to optimise the cell number to be seeded. This is because depletion of BAP1 led to lower number of cells at the end of 72 hours of transfection as evident by the lower protein concentration of lysates obtained (Figure 4.8). The depletion of BAP1 using oligos 1, 2 and 3 led to at least 40% less cells compared to control samples after 72 hours incubation post-transfection, whereas the reduction in cell number for oligo 4 was less dramatic. A rescue experiment involves siRNA transfection and DNA transfection, which cause stress to the cells and can lead to extensive cell death. Therefore, more cells should be seeded at the beginning of a rescue experiment to allow enough proteins to be harvested at the end of the experiment for biochemical analysis.

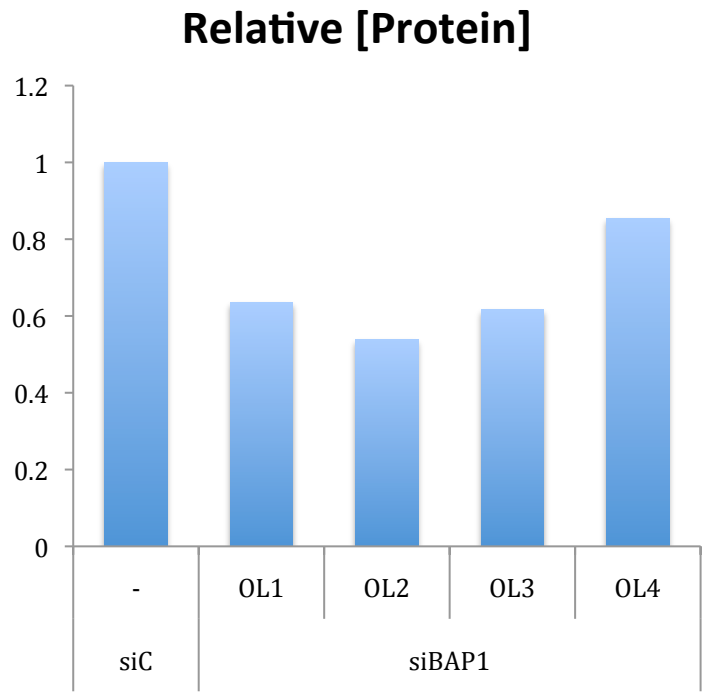

Figure 4.8. Relative protein concentration of MCF7 cell lysates following BAP1 knockdown. MCF7 cells were transfected using non-targeting siRNA oligos or siRNA oligos against BAP1 at 40nM for 72 hours. Cell were lysed using NP40 lysis buffer 72 hours later. Bradford protein assay was performed to determine protein concentration of lysates. Data shown represents average of 2 experiments.

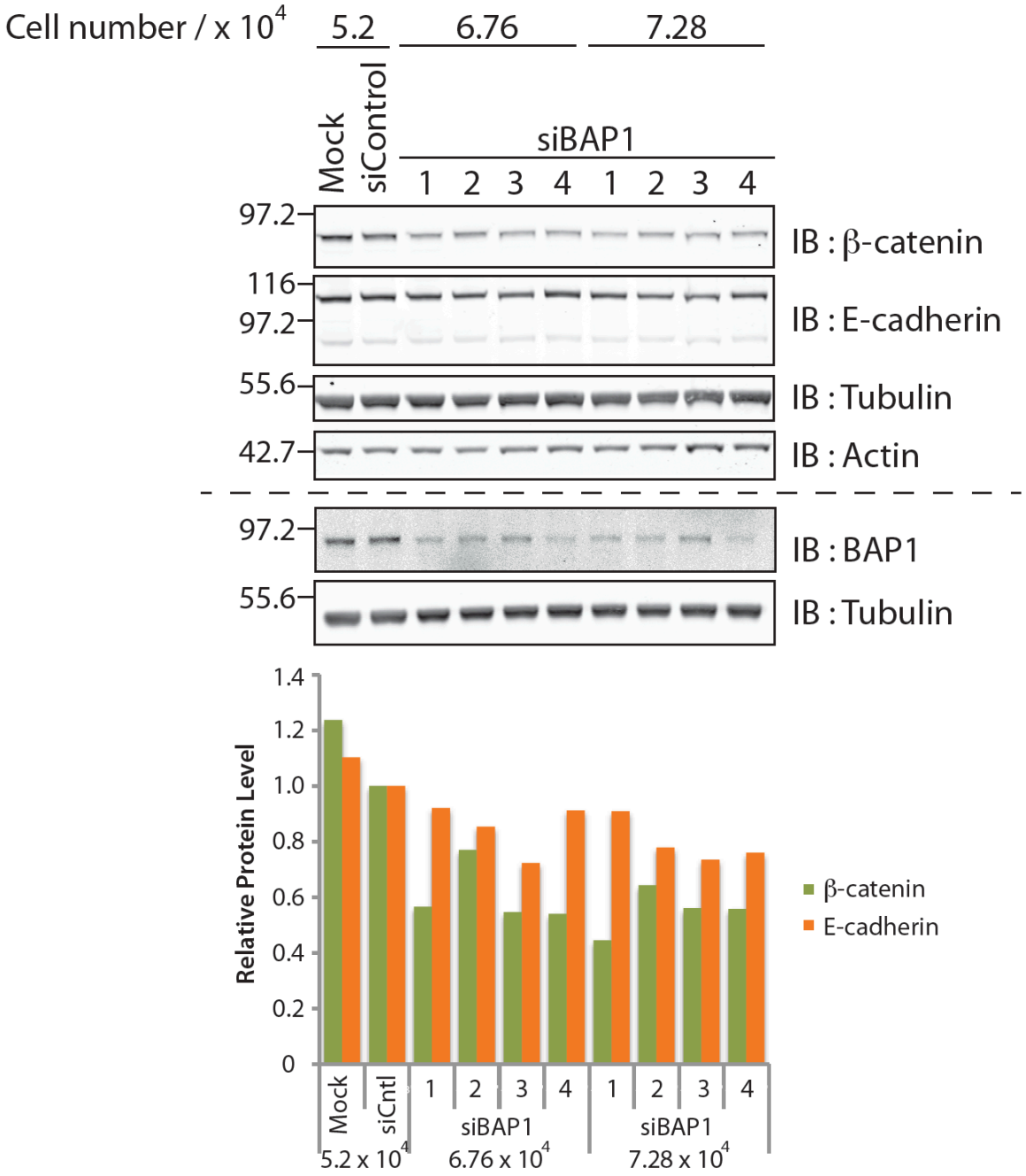

Figure 4.9. The loss of **β–catenin and E-cadherin following BAP1 depletion is not dependent on cell density.** MCF7 cells were transfected using siRNA single oligos against BAP1 at 40nM for 72 hours. Cell were lysed using NP40 lysis buffer 72 hours later. Lysates were resolved on 4-12% NuPAGE gel and immunoblotted for β–catenin, E-cadherin, BAP1, tubulin and actin. Gel images were acquired by infra-red scanner (Odyssey, LICOR). Gel images separated by dotted lines were acquired from different gels.

For this purpose, I have seeded 30% (6.56 x 10^4 cells/well) or 40% (7.28 x 10^4 cells/well) more cells in wells of 12-well plates for transfection using siRNA oligos against BAP1, compared to transfection using non-targeting siRNA (5.2 x 10^4 cells/well). For both cell densities, BAP1 was significantly depleted by all the siRNA oligos (Figure 4.9), and this was accompanied by a reduced level of β-catenin protein in all cases as compared to the control samples (mock and siControl). In this experiment, the level of E-cadherin protein was also quantified. For both cell densities, BAP1 depletion using oligos 1 resulted in a marginal decrease in E-cadherin protein, whereas knockdown using oligos 2 and 3 resulted in reduction of E-cadherin protein level by about 20-30%. For oligo 4, the E-cadherin protein level was similar to that of the control sample for both cell densities.

I have then decided to seed 40% more cells per well (i.e. 1.82 x 10^5 cells/well of a 6-well plate) for the rescue experiment. In order to accumulate more cells expressing BAP1 (wildtype or mutants), I performed the DNA plasmid transfection the day after siRNA transfection, so that the cells had 48 hours (instead of 24 hours) to express BAP1 prior to cell lysis.

I have also tried to generate BAP1 plasmids, which are resistant to siRNA oligos by site directed mutagenesis, but successful mutation was only obtained for resistance against BAP1 siRNA oligo 3. Therefore, I used BAP1 siRNA oligo 3 for BAP1 depletion. The rescue experiment was performed twice and the transfection efficiency of these experiments was between 28-33%. Depletion of BAP1 resulted in lower level of β-catenin and E-cadherin, compared to the control samples, regardless of the plasmid that was transfected. Among the BAP1 depleted samples, the β-catenin level of cells transfected with GFP-BAP1 and GFP-BAP1-C91S was slightly higher than that of cells transfected with GFP, indicating rescue of β-catenin by BAP1, but ,at face value, the rescue is independent of catalytic activity of BAP1. However, the level of β-catenin was only partially recovered as it was still lower than that of the control samples. This is not unexpected since only about one-third of the cells were transfected with the siRNa resistant GFP-BAP1. Expression of GFP-BAP1-A95D on the other hand did not increase the level of β-catenin.

The different GFP-tagged BAP1 proteins showed differential expression level. The wildtype BAP1 always had a higher expression level compared to the BAP1-C91S and BAP1-A95D mutants, and that the expression level of BAP1-A95D mutant was the lowest among the different samples. Intriguingly, the expression level of BAP1-A95D mutant in cells already depleted of endogenous BAP1 was significantly much lower than cells, which were mock-transfected or transfected with non-targeting siRNA oligo. When the blot image for immunoblotting using GFP antibody was adjusted to higher intensity level, smear of higher molecular weight GFP-tagged wildtype BAP1 and BAP1-C91S mutant was observed. While the higher molecular weight smear of BAP1-C91S mutant was more prominent than the wildtype BAP1, the higher molecular weight form of BAP1-A95D mutant was hardly visible.

Multiple studies have identified BAP1 as a tumour suppressor gene (Ventii *et al.*, 2008; Machida *et al.*, 2009), drawing immense interest to understand the biological function of this protein and its relevance to cancer progression. From the experimental results in the previous section, there was clearly a functional link between β-catenin and BAP1, and the different mutations in BAP1 resulted in different fate of the protein itself.

To further characterise BAP1, I have performed an immunoprecipitation experiment by pulling down the GFP-tagged proteins and confirmed that the higher molecular weight smear was polyubiquitylated form of GFP-BAP1. The GFP-BAP1-C91S was much more heavily ubiquitylated compared to the wildtype GFP-BAP1, whereas polyubiquitylation of the mutant GFP-BAP1-A95D was not detected (Figure 4.11). Both the BAP1 mutants and wildtype BAP1 showed nuclear localisation and diffuse localisation in the cytoplasm (Figure 4.12). However, the GFP-BAP1-A95D showed more prominent cytoplasmic localisation than the wildtype BAP1 and BAP1-C91S.

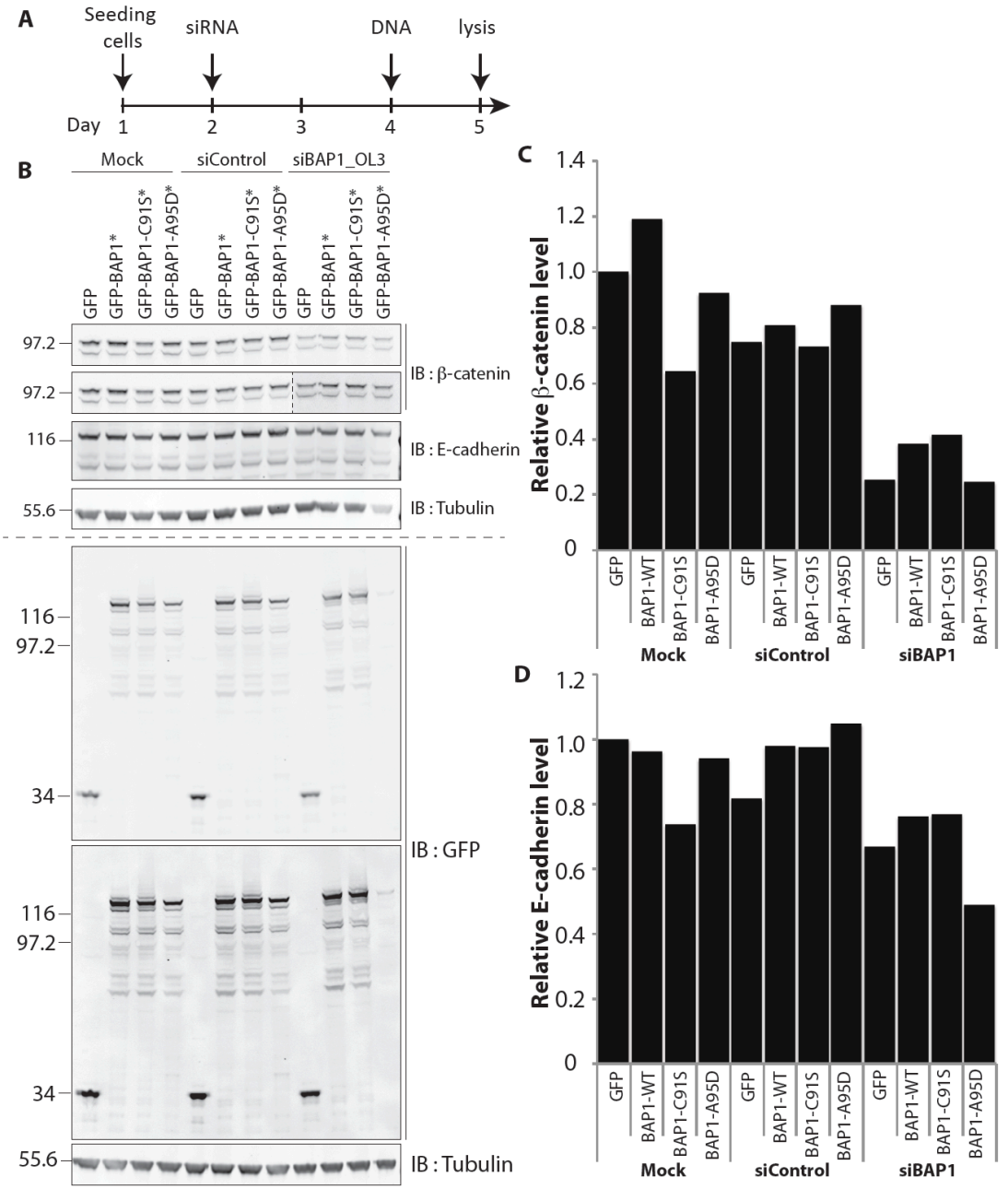

Figure 4.10. Transient expression of siRNA resistant BAP1 rescues β–catenin following loss of endogenous BAP1. (A) MCF7 cells were transfected without siRNA (Mock), or non-targeting siRNA oligo or siRNA oligo 3 against BAP1 at 40nM the day after seeding. 48 hours later, cells were transfected with the different plasmids. Cells were lysed another 24 hours later using NP40 lysis buffer.**(B)** Lysates were resolved on 4-12% NuPAGE gel and immunoblotted for β–catenin, GFP, E-cadherin and tubulin. Blot images were acquired by infra-red scanner (Odyssey, LICOR). Plasmids marked with "*" are siRNA resistant to siBAP1-OL3. Densitrometric analysis of blot was performed to determine relative level of **(C)** β–catenin and **(D)** E-cadherin following normalisation to tubulin.

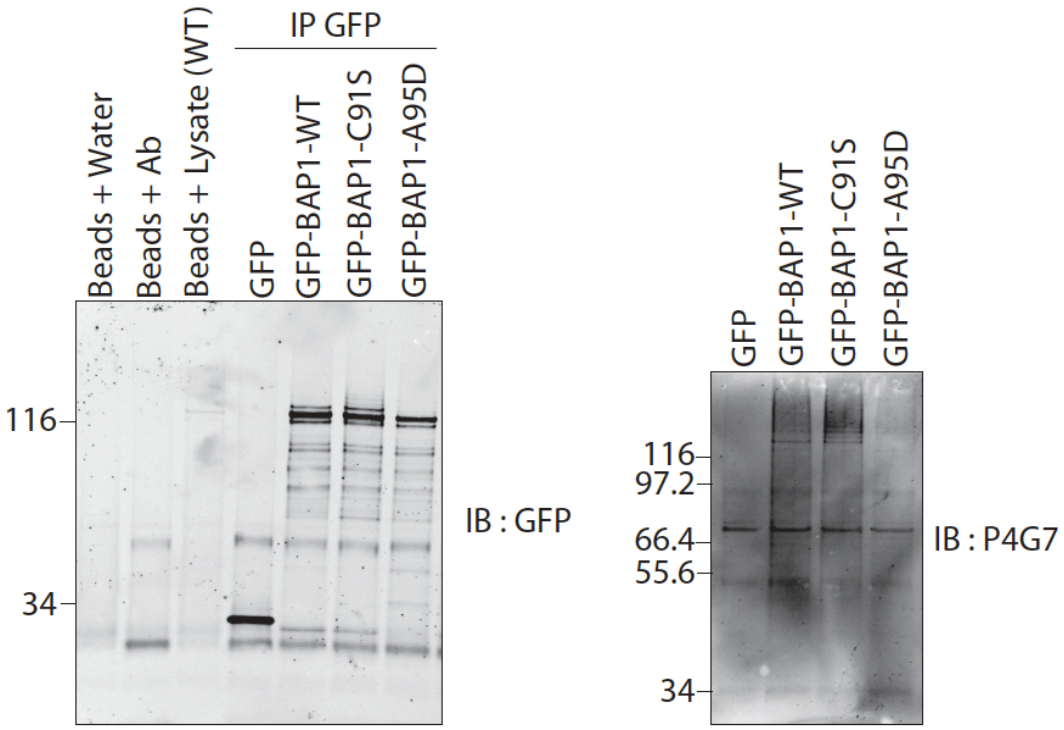

Figure 4.11. Wildtype and C91S mutant BAP1 are polyubiquitylated but not BAP1-A95D mutant. 5.8µg of each of the indicated plasmids was transfected into MCF7 cells grown in 10cm dish and cells were lysed 24 hours later using NP40 lysis buffer. For co-immunoprecipitation, 600µg of protein was incubated with 1µl of GFP antibody and protein G sepharose beads in a final volume of 480µl, for 1.5 hour at 4˚C on a rotating wheel. Purified protein complexes were resolved on 4-12% NuPAGE gel, transferred to nitrocellulose membrane and immunoblotted for GFP or ubiquitin.

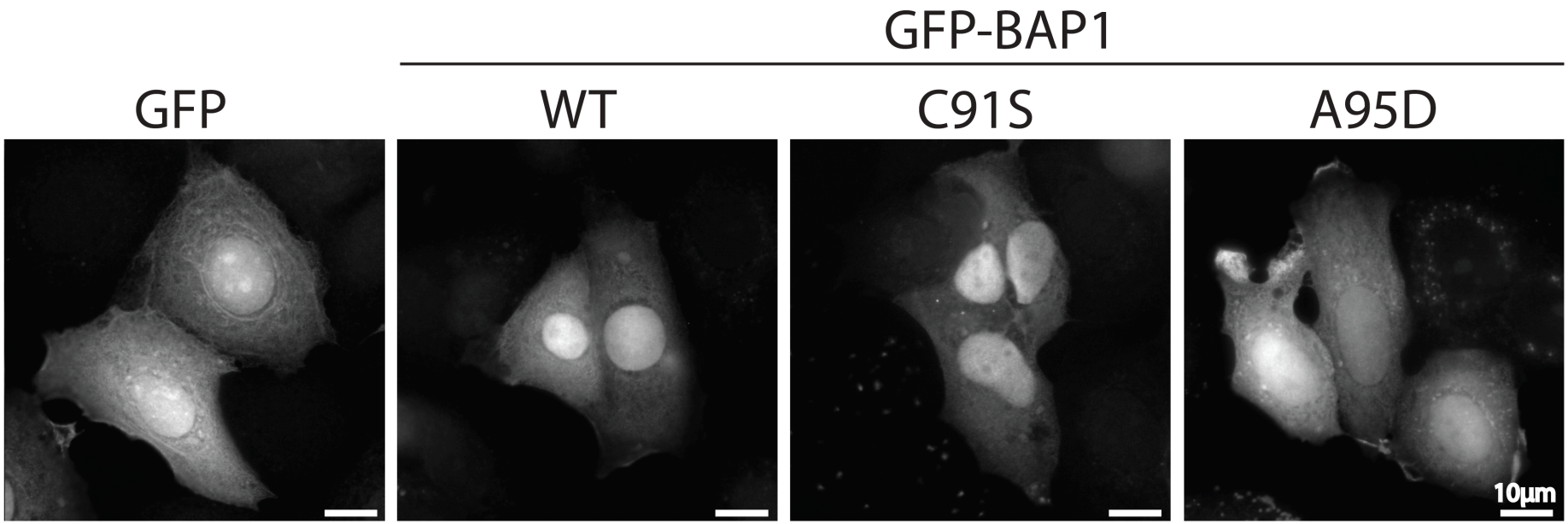

Figure 4.12. Localisation of wildtype and mutant GFP-BAP1. 0.5µg of GFP or 1µg of GFP-BAP1 (WT, C91S and A95D) plasmids were transfected into MCF7 cells. 24 hours later, cells were fixed with 4% PFA/PBS.

4.3 Discussion

By using A549 cell lysates from a large scale DUB human library siRNA screen for Western blotting analyses, I shortlisted 6 DUBs for subsequent verification by deconvolution of the pool of oligos. In the screen, siRNA depletion of TNFAIP3, USP27X and USP54 resulted in at least a two-fold increase in β-catenin level. Notably, the knockdown of USP27X and TNFAIP3 also resulted in significant increase in E-cadherin, which together with β-catenin, is the core component of the E-cadherin/catenin/actin complex at the AJ (Ozawa *et al.*, 1989) and that the efficient coupling of these two proteins is essential for the targeting of the complex to plasma membrane (Chen *et al.*, 1999).

TNFAIP3, also known as A20, was first identified as a tumour necrosis factor alpha (TNFα)-inducible gene product (Dixit, Green *et al.* 1990; Opipari, Boguski *et al.* 1990). Whilst significant work has been done to elucidate the mechanism by which this protein down-regulates NF-κB signaling (Song *et al.*, 1996; Lee *et al.*, 2000; Wertz *et al.*, 2004), there is no previous report of a direct link between TNFAIP3 with regulation of cell adhesion or components of the AJ. The biological functions of USP27X and USP54 are still yet to be elucidated. In other words, there was no data to support a possible role of these DUBs in regulation of AJ components. Using a criterion of at least 2 oligos recapitulating pool knockdown effect, I concluded that the observed increase in β-catenin following siRNA depletion of these DUBs in the screen could be due to off-target effects and these targets were not pursued further.

On the other hand, the siRNA depletion of BAP1, USP9Y and DUB4 resulted in significant decrease in β-catenin. Among these, the decrease in β-catenin following siRNA depletion of USP9Y and DUB4 was potentially a result of off-target effect as well.

Among all the candidates that were deconvoluted, BAP1 emerged as the leading candidate DUB, for involvement in β-catenin regulation and this role of BAP1

was confirmed in a lung cancer cell line, A549 and most strikingly in a breast cancer cell line, MCF7. At least in MCF7 cells, BAP1 is acting as a positive regulator of β-catenin, as its depletion led to the decrease of β-catenin protein level (Figure 4.4A) whereas its overexpression led to an increase in β-catenin protein level (Figure 4.7A). However, the effect of BAP1 on β-catenin was not confirmed in SW480 cells (Figure 4.4B). Although all siRNA oligos against BAP1 resulted in efficient silencing of BAP1, only oligo 1 led to a significant decrease in β-catenin level. One potential explanation to this is that SW480 cell is expressing an isoform of BAP1, which is not targeted by the other oligos. However, among the predicted splice variants of BAP1 on Aceview, none of them is uniquely targeted by oligo 1 only but not the other oligos. An alternative method to confirm the role of BAP1 in this case is to reintroduce a siRNA resistant (to oligo 1) version of BAP1 into SW480.

β-catenin is known to bind to the cytoplasmic domain of E-cadherin (Ozawa *et al.*, 1989) and the association of these two proteins is required for the efficient targeting of the E-cadherin/β-catenin complex to the plasma membrane (Chen *et al.*, 1999). It has also been demonstrated in a human gastric cancer cell line, HSC-39, that cell-cell adhesion was impaired due to mutation in β-catenin, which prevented formation of stable AJ, and hence cell aggregation and cell compaction, despite high level of expression of E-cadherin (Kawaniski *et al.*, 1995). Therefore, the partial loss of cell-to-cell contact and less compact monolayer of epithelial cells following siRNA depletion of BAP1, was likely a secondary effect of the decrease in β-catenin level. Moreover, E-cadherin level was also decreased following BAP1 knockdown (Figure 4.10).

I demonstrated that the knockdown of BAP1 led to a decrease in β-catenin mRNA level in MCF7 cell (Figure 4.6). A similar observation was made by Harbour *et al.*, where siRNA depletion of BAP1 in a uveal melanoma cell line resulted in change in expression levels of a panel of genes, among which included β-catenin and it was downregulated by about 20% upon BAP1 knockdown (Harbour *et al.*, 2010). Moreover, at least in HeLa cells, BAP1 was found to be predominantly complexed

with host cell factor 1 (HCF1), which is a chromatin-associated protein and Yin Yang 1 (YY1), a transcription factor (Machida *et al.*, 2009). The catalytic activity of BAP1 was important to regulate the activity of this whole complex in controlling expression of genes involved in regulation of cell growth and proliferation (Yu *et al.*, 2010). Together, these observations suggest that BAP1 regulates β-catenin at the transcriptional level.

The rescue experiment also indicated a regulatory role of BAP1 on β-catenin (Figure 4.10). In BAP1 depleted cells, the overexpression of the siRNA resistant wildtype BAP1 resulted in about 50% rescue of β-catenin relative to the cells overexpressing GFP. Considering the transfection efficiency of cells was between 28% and 33%, the rescue effect was very prominent. In the series of rescue experiments, I have also included two mutants of BAP1, namely BAP1-C91S and BAP1-A95D. Cys91 of BAP1 is the conserved cysteine residue at the catalytic triad responsible for cleavage of isopeptide bond and Ala95 is a well conserved amino acid residue within the UCH domain, close to the active site. Both of the mutants have been shown to lack deubiquitylating activity based on an *in vitro* deubiquitylation assay using ubiquitin COOH-terminal 7-amido-4-methylcoumarin (Ub-AMC) as substrate (Ventii *et al.*, 2008). However, the two mutants did not lead to the same effect on β-catenin.

The BAP1-C91S mutant when overexpressed in cells with endogenous level of BAP1 did not alter the level of β-catenin, unlike the overexpression of wildtype BAP1. This suggests that the positive regulatory role of BAP1 on β-catenin depends on its catalytic activity (Figure 4.7). However, contradictory to this observation, the BAP1-C91S mutant was able to recover β-catenin in BAP1 depleted cells, to a similar extent as the wildtype BAP1, suggesting that the catalytic activity of BAP1 is dispensable for its regulatory role on β-catenin expression. As mentioned before, BAP1 was found in complex with HCF1 and YY1, and the multiprotein complex plays a role in regulating expression of a wide array of genes. While the catalytic activity of BAP1 is dispensable for the assembly of the complex (Yu *et al.*, 2010), work by

Machida and colleagues suggests that the DUB activity is required to deubiquitylate HCF1 to allow regulation of cell proliferation (Machida *et al.*, 2009). They also demonstrated that most, if not all, endogenous BAP1 is found in complex with HCF-1. HCF-1 was also found to be the top interacting partner of BAP1 in mouse as identified by mass spectrometry (Dey *et al.*, 2012).

In light of the above evidence, it is very likely that β-catenin expression is also regulated by the transcriptional complex containing BAP1 and HCF-1. There is one possible explanation for the apparent rescue of β-catenin by the catalytically inactive BAP1-C91S mutant: in a normal cell, there are multiple BAP1-containing transcription activating complexes, shifting in equilibrium between inactive (HCF1 ubiquitylated) or active (HCF1 not ubiquitylated) state (Figure 4.13A). Therefore, when wildtype BAP1 is overexpressed, there are more active transcription activating complexes due to enhanced deubiquitylation of HCF1 by BAP1, resulting in a higher expression level of β-catenin (Figure 4.13B); whereas overexpression of the catalytically inactive mutant C91S does not alter the active-inactive equilibrium of the transcription complexes (Figure 4.13C). However, in cells where endogenous BAP1 is largely depleted, the exogenous BAP1, whether wildtype or mutant, is incorporated into the transcription complexes. These complexes are mostly retained in the active state (i.e. not ubiquitylated), due to limiting level of BAP1, to maintain regulation of gene expression, which means the deubiquitylating activity of BAP1 is dispensable (Figure 4.13D). Therefore, in cells where BAP1 was already depleted, both the wildtype and C91S mutant of BAP1 can rescue β-catenin (Figure 4.13E & F).

The overexpression of the BAP1-A95D mutant, which also lacks deubiquitylating activity, did not rescue β-catenin level. If the above mentioned mechanism is true, this observation is not surprising since the expression level of this mutant was extremely low, as compared to the other BAP1 proteins. In that case, the formation of BAP1-HCF-1 transcriptional activating complex was limited to allow efficient rescue of β-catenin. However, it is not known and not determined during my research project whether the A95D mutant can interact with HCF1 or not.

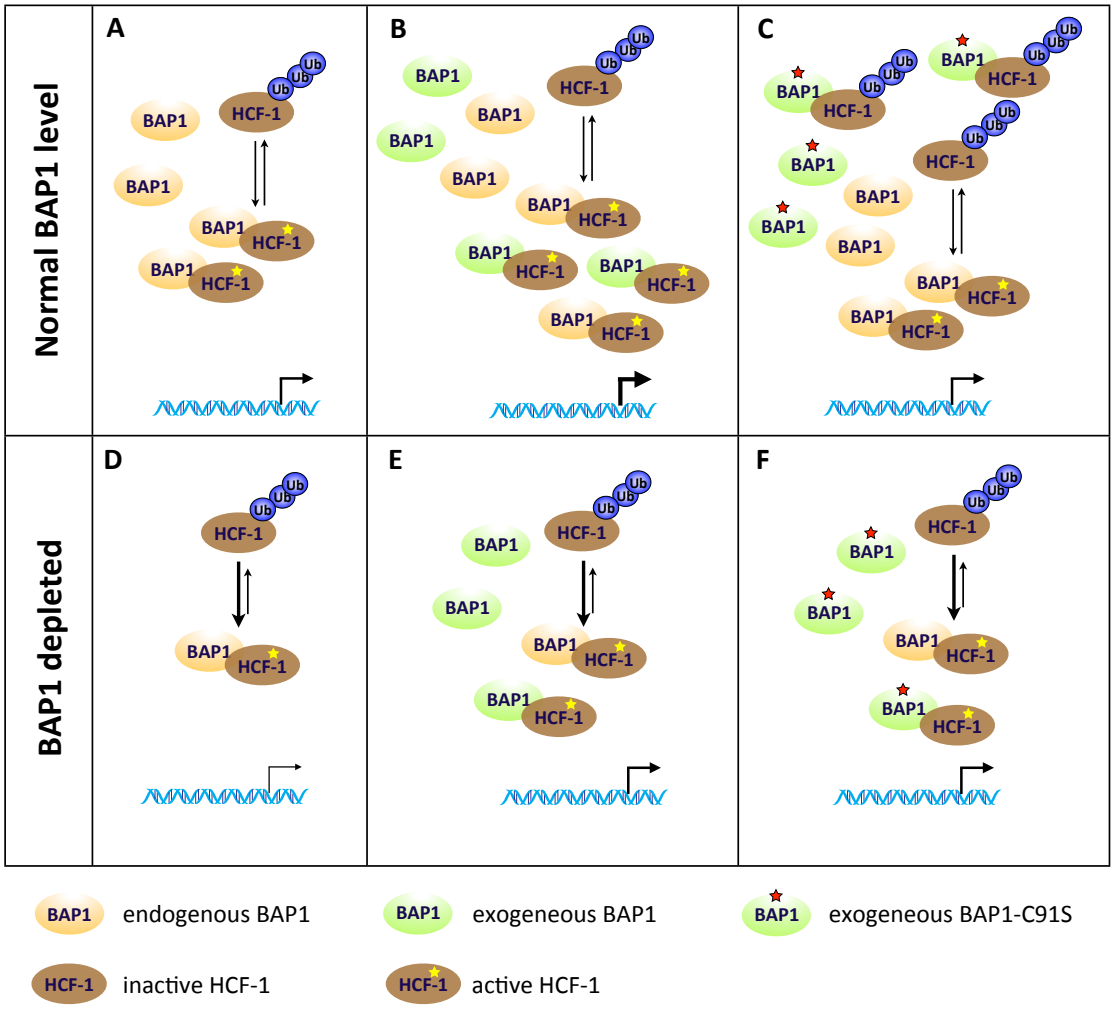

Figure 4.13. Possible mechanism of β-catenin transcriptional regulation by the BAP1-HCF-1 complex. At steady state, **(A)** HCF-1 constantly shifts in equilibrium between an activated deubiquitylated state and an inactive polyubiquitylated state. Association of HCF-1 with BAP1 leads to its deubiquitylation and hence activation to enable gene expression. **(B)** When wildtype BAP1 is transiently overexpressed, more active complexes of BAP1-HCF-1 are formed, thereby increasing β-catenin expression level whereas **(C)** overexpression of BAP1-C91S mutant does not shift the equilibrium and hence β-catenin expression level remains unaltered. **(D)** When BAP1 is depleted, the active-inactive HCF-1 equilibrium favours a non-ubiquitylated form of HCF-1, to maintain basal gene expression. Therefore, whether **(E)** wildtype or **(F)** C91S mutant of BAP1 is overexpressed, both of them are recruited to form an active transcriptional complex.

The rescue experiment also yielded some interesting observations, highlighting the different biological outcome resulting from different amino acid mutations in BAP1. Firstly, the C91S mutant is more heavily polyubiquitylated

compared to wildtype BAP1, which is an observation that has already been reported (Yu *et al.*, 2010). This is also a common observation among catalytically inactive mutants of DUBs, such as USP4 (Wada and Kamitani, 2006) and USP8 (Mizuno *et al.*, 2005), which are unable to self-deubiquitylate. However, such polyubiquitylation of A95D mutant was not observed. It is puzzling how the mutation of a non-lysine residue would abolish ubiquitylation of a protein. A structure of the mutant may provide insight into the loss of ubiquitylation, as a single amino acid substitution may result in dramatic conformational change. For example, it may abolishes the ability of BAP1 to bind to an E3 ligase. Understanding the distinct biology of the A95D mutant will certainly provide further insights into the cellular roles of BAP1. In the same light of thought, it would be interesting to investigate the biological consequence of other BAP1 mutations, such as Q36X, G128R, H169Q and S172X in the UCH domain, and other mutations which are found along the C-terminal extension of BAP1 (Harbour *et al.*, 2010).

Secondly, the expression level of the A95D mutant is always lower than the wildtype BAP1 and C91S mutant in all conditions. More intriguingly, the expression level of the C91S mutant was similar in cells whether or not BAP1 is depleted, whereas the A95D mutant was not well expressed in cells already depleted of BAP1. This indicates that the presence of endogenous BAP1 is required to maintain protein stability of the A95D mutant. However, the former scenario is less likely, because if the endogenous BAP1 stabilises the A95D mutant by deubiquitylating it, polyubiquitylation of the A95D mutant would be expected, but that was not the case (Figure 4.11). If the low expression level of the A95D mutant (which is not polyubiquitylated like the wildtype BAP1 and C91S mutant) is due to a transcriptional defect, it is tempting to speculate that ubiquitylation of BAP1 itself is required for the positive modulation of its transcriptional activity or that BAP1 ubiquitylation is essential for its incorporation into a transcriptional complex. ATXN3 is the first reported DUB for which its ubiquitylation enhances its own catalytic activity (Todi *et al.*, 2009). In addition, the different sub-cellular distribution of the A95D mutant may account for its distinct biology as well.

In summary, the work presented in this chapter has revealed the functional relationship between two proteins which are highly relevant to cancer, namely BAP1, a tumour suppressor gene which has been found to be somatically and germline mutated (see section 1.2.3.6.1) (White and Harper, 2012) and β-catenin, a proto-oncogenic gene whose aberrant stabilisation leads to activation of the Wnt signaling pathway, which is associated with tumorigenic phenotypes such as enhanced cell proliferation and acquisition of metastatic features (see section 1.1.2.2.3) (Reviewed in Giles *et al.*, 2003; Clevers, 2004; MacDonald et al., 2009). However, the mechanism by which the regulatory role of BAP1 on β-catenin, at the transcriptional level, is exerted is not fully elucidated. Much work has been done to understand how β-catenin is regulated post-translationally but our understanding of the transcriptional regulation of this protein is very limited. Therefore, my finding here provides an important insight into the transcriptional regulation of β-catenin, through the involvement of a deubiquitinase, BAP1. Of note, BAP1 is regulating β-catenin gene expression at steady state conditions, and the effect of BAP1 depletion or overexpression on transduction of Wnt signaling has not been assessed. Taking the face value of the data presented in this chapter, it would be anticipated that BAP1 plays a positive regulatory role in Wnt signaling, implying an oncogenic property of BAP1. Accordingly, the inhibition of BAP1 in the treatment of cancers arising from aberrant activation of Wnt signaling is a reasonable and potentially promising therapeutic intervention. However, such strategy would be a contradiction to the clinical and experimental data collated so far, which demonstrate a tumour suppressive role of BAP1 and its inhibition may have tumorigenic effect. This highlights the need to understand the molecular events that underlie the BAP1-β-catenin in the context of Wnt signaling and cancer.

Chapter 5

Construction and Characterisation of a Human Deubiquitylase esiRNA Library

5.1 Introduction

An alternative siRNA screening strategy is to use endonuclease-prepared siRNA (esiRNA), which potentially offers solution to issues with siRNA. Quite regularly, some of the potential candidates identified in siRNA screenings fail the subsequent validation phase of deconvolution experiments due to off-target effects, a common feature of siRNA. In other words, there is always an inevitable waste of money, effort and time associated with siRNA screening. Therefore, we decided to generate a human DUB endonuclease-prepared siRNA (esiRNA) library, to (i) reduce incidents of false positive results due to off-target effects, (ii) reduce cost, (iii) complement current siRNA libraries as an alternative option for validation and (iv) serve as a resource for the ubiquitin community. To verify the utility of this resource, the DUB esiRNA library generated requires diligent characterization in term of

knockdown efficiency of individual esiRNAs, consistency of silencing effects, knockdown efficiency across a variety of cell lines and toxicity to cells.

5.1.1 Endonuclease-prepared siRNA (esiRNA)

In 2002, a different flavour of siRNA, generated by enzymatic digestion of double stranded-RNA (dsRNA), also known as endonuclease-prepared siRNA (esiRNA) was introduced by Michael Bishop's group and this technology was later pioneered by Frank Buccholz (Yang et al., 2002). Although considered a novel technology, the concept of esiRNA is not new since the mechanism of silencing is the same as siRNA. As Frank Buccholz said, "esiRNA technology was developed at the same time RNAi hit mammalian cells". Basically, generation of esiRNA is simply mimicking the in vivo digestion of dsRNA in C. elegans in vitro with the help of RNAse III enzyme from Escherichia coli instead of the eukaryotic Dicer. esiRNA is a very attractive alternative to chemically synthesized siRNA as it is easy to generate and available at relatively low price, without compromising the silencing efficiency. It is also suggested that if offers a greater coverage of a larger portion of endogenous mRNA, hence higher penetrance of the knockdown effect. More importantly, it can potentially circumvent the problem of off-target effects and variability of chemically synthesized siRNA, due to the high complexity of the pool of siRNAs. Several large scale RNAi screening experiments have been performed using esiRNA, yielding insightful information and uncovering promising targets for the understanding of cellular processes such as cytokinesis (Zhu and Jiang, 2005; Zhu *et al.*, 2005) and cell cycle (Kittler *et al.*, 2004; Kittler and Buccholz, 2005).

5.2 Generating a DUB esiRNA Library

In brief, the production of esiRNA involves PCR amplification of a specific gene sequence followed by *in vitro* transcription to generate dsRNA which is then

enzymatically digested to produce a pool of esiRNA, 20-25 nucleotides in length (Figure 5.1).

5.2.1 Design of esiRNA Primers

5.2.1.1 DEQOR

The very first step towards generating an esiRNA library is to select the gene sequence from which the corresponding esiRNA is to be generated and hence the required primer pairs complementary to the sequence at the boundary can be designed . For this purpose, a web-based tool, named DEQOR, was developed for the design and quality control of esiRNAs (Figure 5.2; Henschel et al., 2004). The first step in a DEQOR analysis is a BLASTN search for the source sequence, so that it will be excluded from subsequent analysis for cross-silencing activity of the corresponding esiRNA. Subsequently, an *in silico* digestion of the input sequence is performed, whereby a step-size of 1-nucleotide is used to slide along the entire length of input sequence and the user can customize the window size for length of siRNA to be generated, between 16 to 25 nucleotides.

In the next step, the *in silico* generated pools of siRNA are analysed for quality control purposes according to optimal siRNA properties. There are three main criteria a siRNA should meet for potent induction of gene silencing: (i) asymmetry on the 5' and 3' ends, with an adenine (A) or thymine (T) at its 5' end and a guanine (G) or cytosine (C) at the 3' end; (ii) no more than 3 tandem repeats of the same nucleotide present within the primer sequence; (iii) the GC content of the target sequence should be within the range of 20-50%. DEQOR utilizes a simple penalty scheme according to the 3 criteria to assess the quality of esiRNA generated: (i) if the siRNA has a "reverse symmetry", i.e. a G/C at its 5' end and an A/T at its 3' end, there is a penalty of 7 points and a penalty of 3 points is imposed if a G/C or A/T is present on both ends; (ii) a 7 point penalty is imposed for the presence of 3 tandem repeats of the same nucleotides; (iii) if GC content of the siRNA is out of the

optimal range, the siRNA is penalized with 1 point for every 1% deviation below 20% GC-content or 1 point for every 2% deviation above 50%.

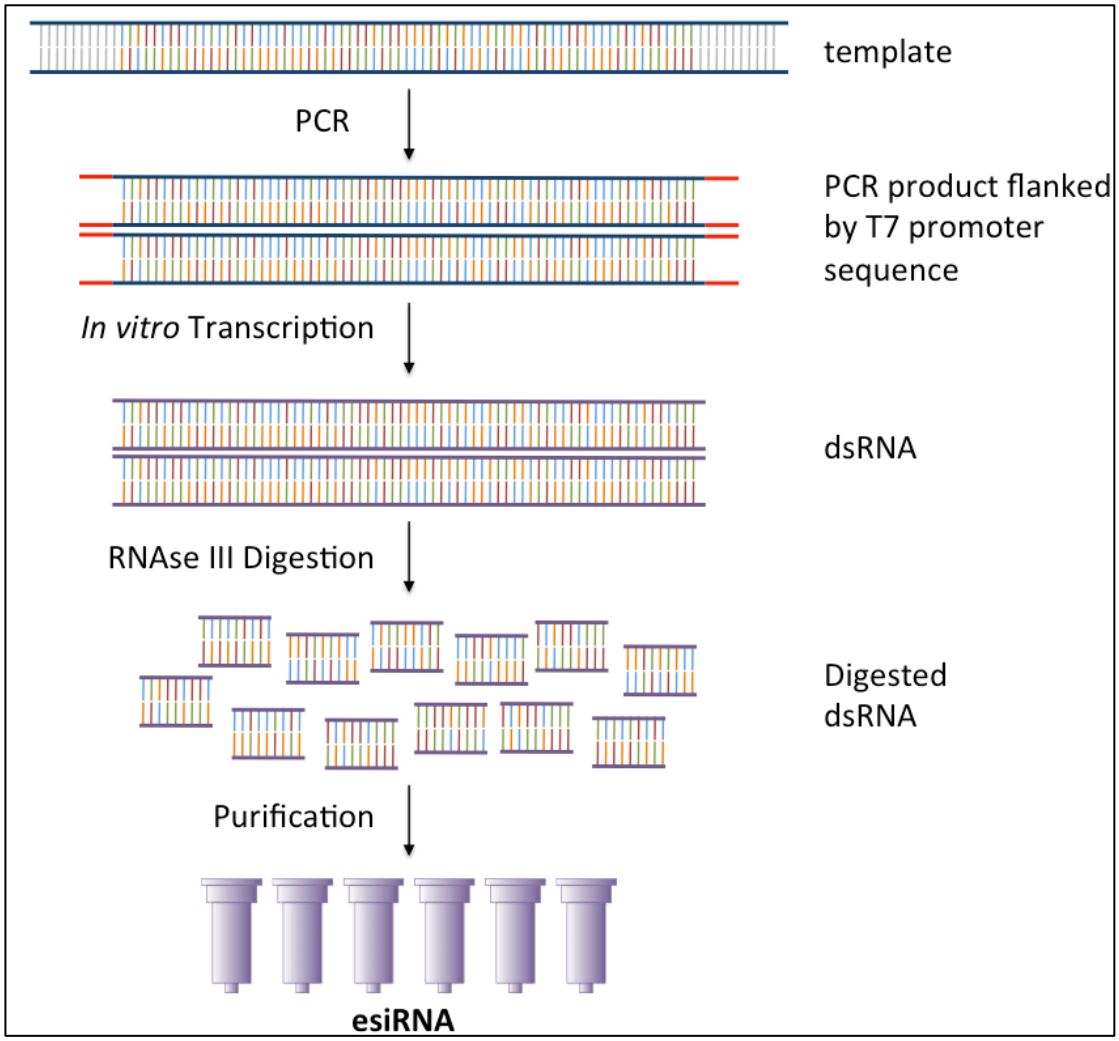

Figure 5.1. **Workflow of esiRNA production.** Selected esiRNA region is amplified using specific primers flanked with the T7 promoter sequence, which allows transcription initiation by T7 RNA polymerase. Purified PCR product is then subjected to *in vitro* transcription to generate double stranded RNA, which is subsequently digested by RNAse III to produce random fragments of dsRNA of 20-25 nucleotides in length. esiRNA is purified by ion-exchange column and resuspended in RNAse-free water.

In addition to their inherent chemical properties, siRNAs are also assessed on their cross-silencing activity. A penalty of 10 points is imposed for perfect match with a non-target mRNA sequence, and for the case of a single mismatch, the siRNA is penalized 8 points for a de-central mismatch and 2 points for a central mismatch.

The default penalty setting of DEQOR is summarized in Table 1. The penalties of all siRNA within the esiRNA mixture are summed up to give a quality score, and the individual siRNAs are ranked according to their scores and siRNAs with score within the range of -5 to 0 are considered to have optimal silencing capacity.

Table 5.1.

	Penalty
Sequence quality	
Reverse asymmetry	7
Symmetry	3
Polynucleotide stretch	7
GC Content	
>50%	1 per 2% deviation
<20%	1 per 1% deviation
Cross-silencing activity	
Perfect match	10
One mismatch	
Central	2
De-Central	8

Table 5.1. Default penalty setting of DEQOR (Henschel *et al.,* 2004)

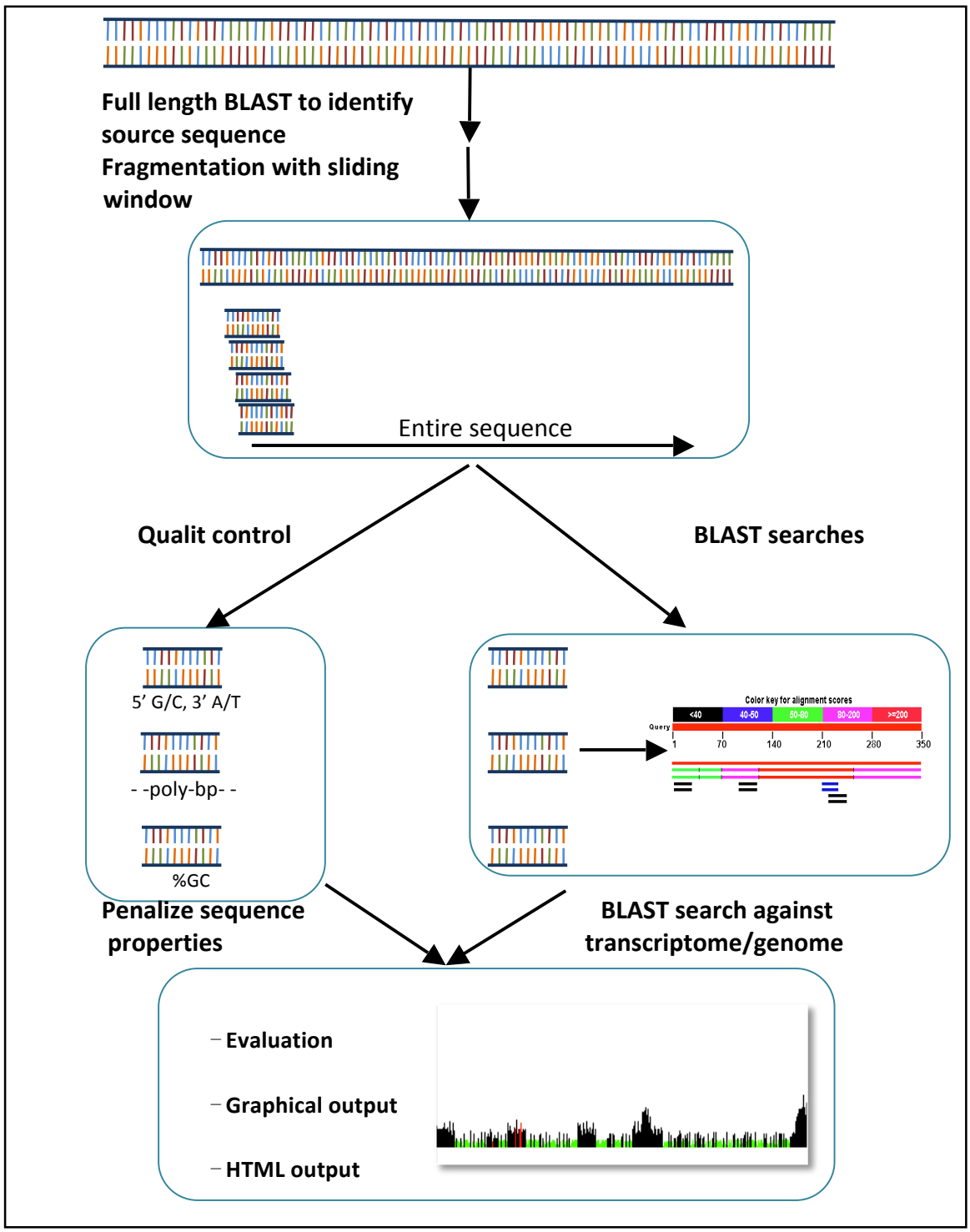

Figure 5.2. **Workflow of a DEQOR analysis.** First, a BLAST search is performed for the input query to identify its origin, followed by *in silico* digestion of the sequence into fragments of 16 to 25 nucleotides which act as siRNA. Each of these siRNAs is assessed for its chemical properties according to the criteria for optimal siRNA and is penalised for any deviation from such standards. In parallel, the siRNAs are subjected to BLASTN search to assess for cross-silencing activity. Together, these analyses allow evaluation of the input query for esiRNA generation and this is summarised by a graphical output. A HTML output is also generated for a list of off-targets.

5.2.1.2 RiDDLE

Frank Buccholz and his group, by using DEQOR, have created a web-based database named RiDDLE, to encourage the use of esiRNA technology(Figure 5.3; Kittler et al., 2007). RiDDLE serves as an online database for design of esiRNA, whereby users are allowed to search for esiRNA for their genes of interest based on identity search or annotation based search.

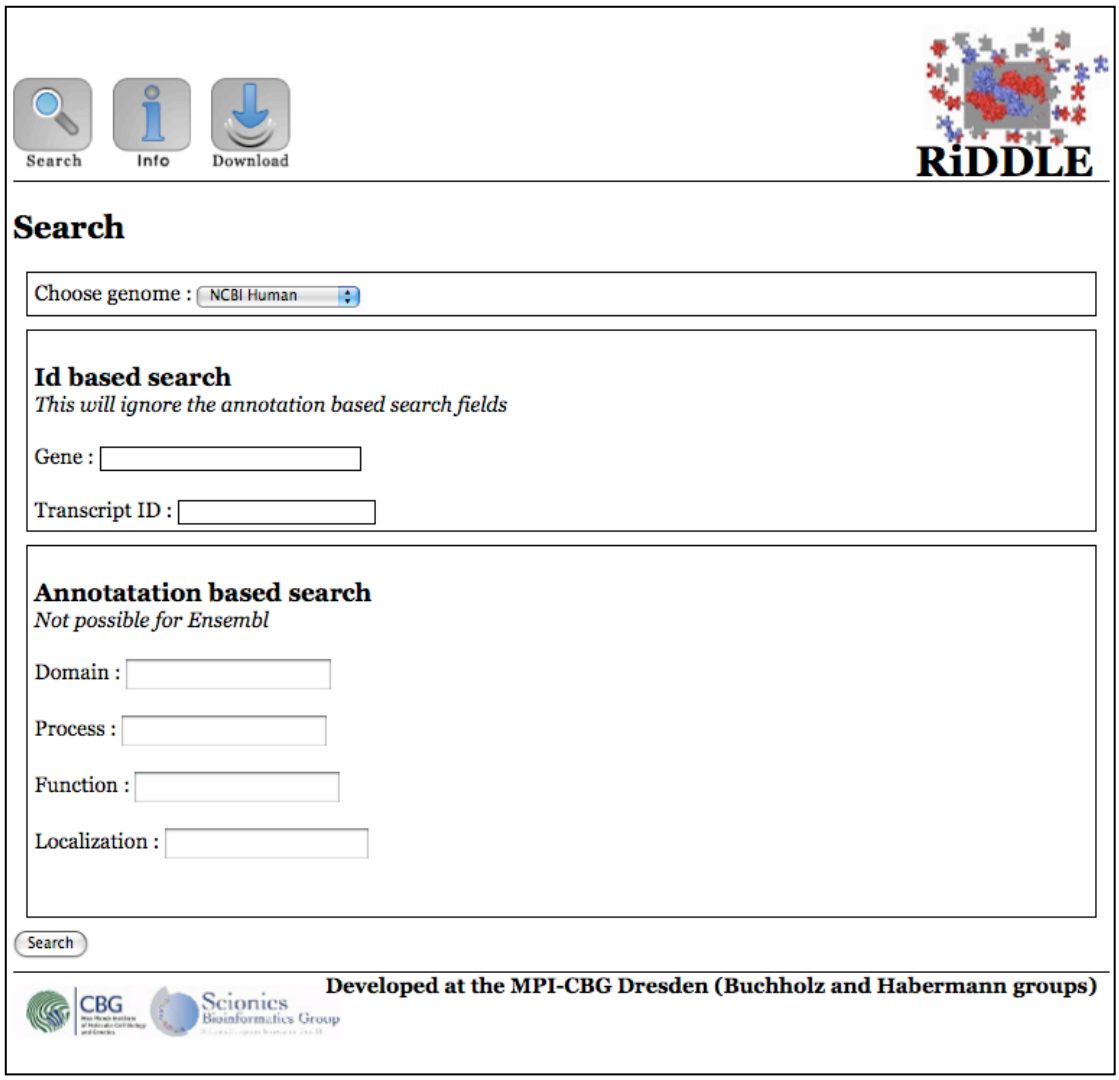

Figure 5.3. **Search page of RiDDLE.** Users are allowed to search for targets of a particular genome, using an ID based search or an annotation based search.

Various pieces of information are provided by a typical search resulton RiDDLE, including esiRNA sequence, primers that amplify that region, efficiency of the predicted esiRNA and annotation based on Gene Ontology (GO). Besides, there

are links, which lead users to a BLAST search ("NCBI", "Ensembl" and "UCSC") and other pages containing relevant information to that particular gene, including paralogues in other species, homologues within the same species and genes containing the same domain as the gene of interest (Figure 5.4).

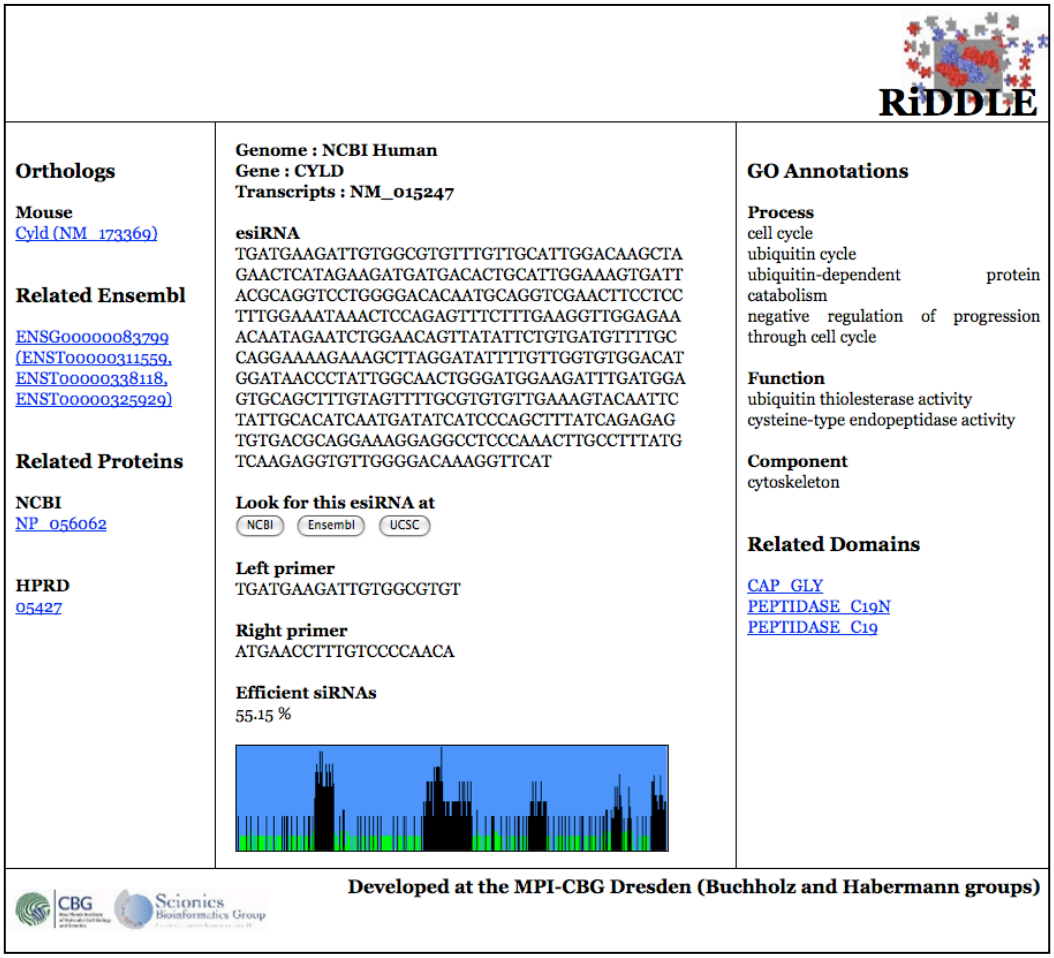

Figure 5.4. Search result for CYLD on RiDDLE.

To begin with primer design, I have first acquired the esiRNA sequence from RiDDLE and the information is available for 90 out of the 93 human DUBs (ATXN3L, USP27X and USP17 are not found on RiDDLE). As seen in Figure 5.4, the primer pairs which amplify the esiRNA region are highlighted.

5.2.1.3 Optimisation of Primers

Although esiRNA regions predicted on RiDDLE are meant to be optimal, there were various considerations that I had to take into account before I decided the primers to be used: (i) template for PCR amplification of esiRNA region; (ii) primer properties; (iii) quality control.

(i) Template for PCR amplification of esiRNA region

Sebastian Hayes, a previous PhD student in the lab has built a collection of plasmids containing the open reading frame (ORF) of 75 DUBs. It was desirable to use these plasmids as template for PCR amplification of esiRNA regions for three reasons: (i) amplification from cDNA can sometimes be challenging due to the high complexity of DNA content and limitation of sensitivity of primers to amplify the correct sequence; (ii) for amplification from cDNA, the correct cDNA library has to be used as not all genes are ubiquitously expressed; (iii) the plasmids have all been sequence-verified and that gives almost absolute confidence of the identity of PCR product purified.

Therefore, I performed alignment analysis for the predicted esiRNA sequence extracted from RiDDLE and the cloned ORFs. Among the 75 alignments done, 53 of them had perfect sequence match, 10 of them had mismatch within the range of 1 to 16 base pairs, 3 of them were predicted using different transcripts, 5 of them extended into either the 5’ or 3’ untranslated region (UTR) of the mRNA, 3 were completely from the UTR and 1 was not found on RiDDLE database.

For DUBs whichwe did not already possess as ORFs, the esiRNA regions were amplified from cDNA.

(ii) Primer properties

To achieve optimal PCR amplification, the physical properties of primers were optimized according to the following criteria:

(a) complementary sequence of at least 18-20 nucleotides;

(b) GC content within the range of 50% to 70%;

(c) no complementary sequence within primers themselves;

(d) no more than 3 tandem repeats of the same nucleotides.

For suggested primers which do not conform to this criteria, I redesigned the primers based on sequence around the boundary of the predicted optimal esiRNA region, thereby minimizing compromises to the esiRNA quality.

(iii) Quality control

(a) Cross-silencing activity and sequence coverage

Although DEQOR is supposed to predict the esiRNA region of optimal quality, I double-checked the cross-silencing activity on other DUBs, coverage for different transcript variants of the same gene and efficiency of the predicted esiRNA region. I performed a BLAST analysis for all the esiRNA sequence, and found 9 of them to have cross-silencing activity on other DUBs (USP6, USP9X, USP9Y, USP18, USP32, USP41, DUB3 and DUB4) and at least 2 of them (USP2A and UCHL5) did not target all the known transcript variants.

(b) Efficiency

While most esiRNAs have predicted efficiency of 50 – 60%, I set a cut off value of 40% and any esiRNA with predicted efficiency of less than 40% is considered to be unsatisfactory. With this criterion, I have shortlisted 7 esiRNAs (USP2A, USP20, USP49, JOSD2, UCHL1, MPND, OTUD6A) with predicted low knockdown efficiency and tried to select for a different esiRNA region, which should give higher esiRNA quality (>40%) and that was possible only for 3 (USP2A, USP49, UCHL1) of the 7.

For primers which do not qualify for any of the above criteria, I redesigned the primers by using DEQOR. Table 5.2 summarises the primer sequences that were ordered for generation of the DUB esiRNA library.

5.2.2 Cloning of esiRNA Region

For DUBs of which the esiRNA regions were to be amplified from cDNA, the main concern was the identity of the amplified sequence, since cDNA is a highly

complex mixture. The size of PCR product, although indicative, does not give absolute confidence of sequence identity.

To address this issue, we decided to insert the PCR-amplified esiRNA region into bacterial expression vectors and the plasmids purified from bacteria are subjected to sequence verification. Having the esiRNA region cloned into plasmids also obviates amplification from cDNA in the future and the esiRNA sequence can be produced in bulk economically by bacterial amplification.

We have decided to employ blunt-end cloning strategy, using Zero Blunt® TOPO® PCR Cloning Kit (Invitrogen),for the direct insertion of PCR products into pCR®4Blunt-TOPO® vector. In the kit, the plasmid vector is linearised and is covalently linked to the Vaccinia virus topoisomerase I on its 3' ends. This is a DNA-enzyme intermediate linked by a phospho-tyrosyl bond between the 3' phosphate group of the vector and tyrosine residue 274 of the enzyme (Figure 5.5; Shuman, 1991). When mixed with PCR product and the mixture allowed to stand at room temperature, this phospho-tyrosyl bond is attacked by the 5'-hydroxyl group of the PCR product with the formation of a phosphodiester bond between the 3'-phosphate group of the vector and 5'-hydroxyl group of the PCR product, while the enzyme topoisomerase I is released.

There are several advantages of using this cloning strategy:

(i) it is highly efficient and rapid, involving only a one-step reaction of 5 minutes;

(ii) no extra sequence needs to be included in the primer for subsequent cloning;

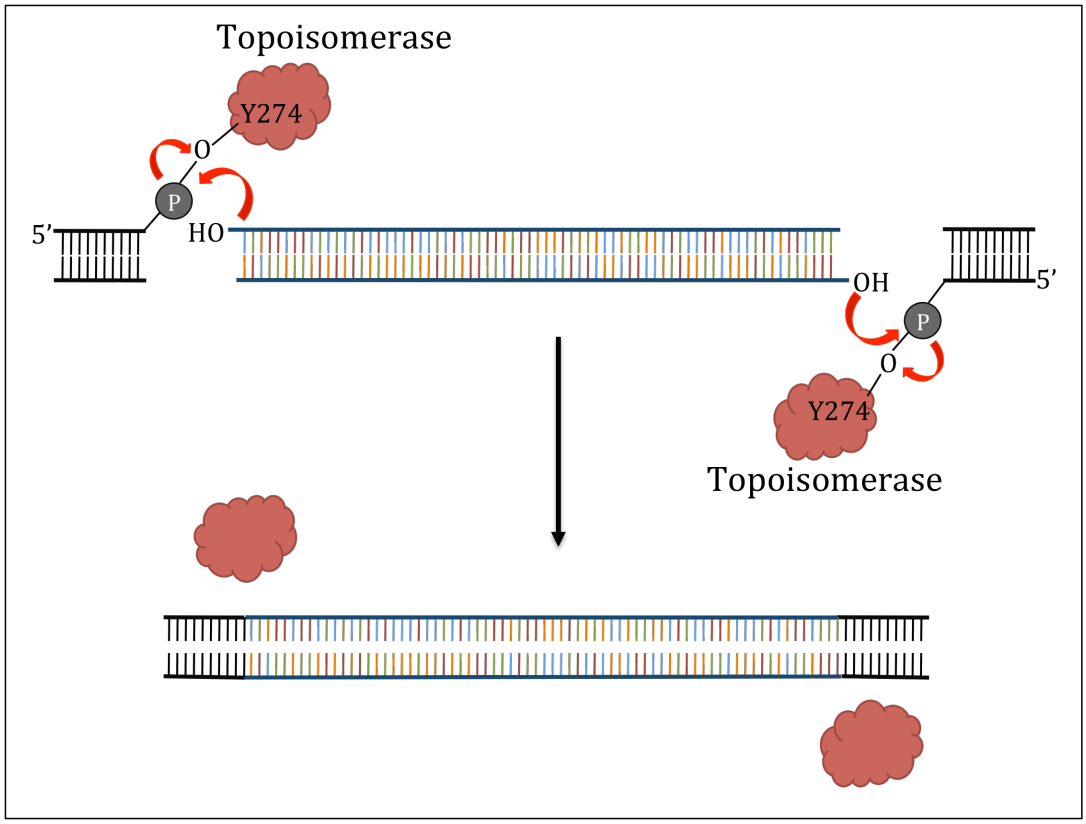

Figure 5.5. **Principle of TOPO-cloning reaction.** The 5' hydroxyl groups of the PCR product attack the phospho-tyrosyl bond between the vector and Topoisomerase, with a concomitant formation of phosphodiester bond between the PCR product and vector.

(iii) the vectors are engineered for sequencing purpose, allowing sequencing using various commercially available primers, such as M13F, M13R and T7.

I have amplified the esiRNA regions from both A549 and MCF7 cDNA. The PCR product purified after DNA agarose gel electrophoresis was then used for Zero Blunt Topo Cloning reaction. 5µl of the reaction mixture was used for transformation of MACH1^{TM} – T1 Phage Resistant E. Coli competent cells and cells were inoculated on Ampicillin agar plates. Colonies which gave positive results for colony PCR on the next day were used to set up 5ml overnight culture for plasmid extraction. Extracted plasmids were subjected to restriction digest before being sent off for sequence verification. A glycerol stock was kept for each clone with positive sequencing results. Table 5.2 summarises the plasmids that were created.

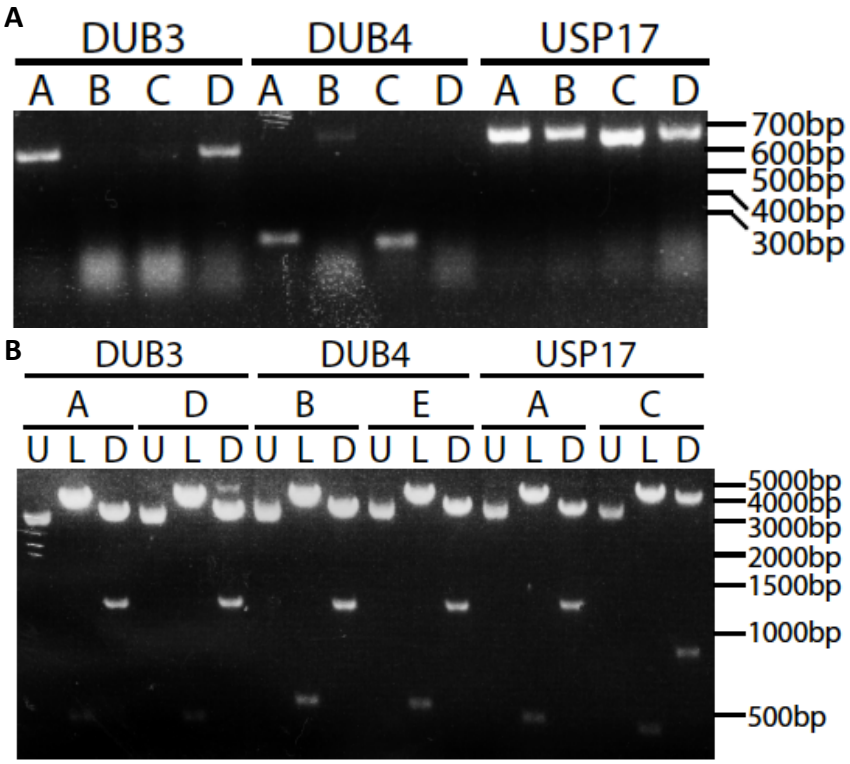

Figure 5.6.**Verification of pCR4Blunt-TOPO plasmids containing esiRNA regions**. **A.** Colony PCR was performed to check presence of esiRNA region. 4 colonies were checked for each bacterial transformation after TOPO cloning reaction. **B.** Plasmids were purified from bacterial colonies qualified for colony PCR and subjected to restriction digest for verification. Note: U – uncut; L – linearised; D – double cut.

Table 5.2.

DUB	Plasmid	Glycerol Stock No.
DUB3	pCR4TOPO-T7-DUB3	JL069
DUB4	pCR4TOPO-T7-DUB4	JL055
USP17	pCR4TOPO-T7-USP17	JL056
USP22	pCR4TOPO-T7-USP22	JL057
USP24	pCR4TOPO-T7-USP24	JL058
USP31	pCR4TOPO-T7-USP31	JL059
USP34	pCR4TOPO-T7-USP34	JL060
USP35	pCR4TOPO-T7-USP35	JL061
USP37	pCR4TOPO-T7-USP37	JL071
USP40	pCR4TOPO-T7-USP40	JL063
USP43	pCR4TOPO-T7-USP43	JL064
USP51	pCR4TOPO-T7-USP51	JL072
ATXN3	pCR4TOPO-T7-ATXN3	JL065
ATXN3L	pCR4TOPO-T7-ATXN3L	JL066
PRPF8	pCR4TOPO-T7-PRPF8	JL070
OTUD1	pCR4TOPO-T7-OTUD1	JL067
OTUD7A	pCR4TOPO-T7-OTUD7A	JL073
PARP11	pCR4TOPO-T7-PARP11	JL068

Table 5.2 pCR4Blunt-TOPO plasmids containing esiRNA regions of DUBs.

5.2.3 esiRNA Generation

5.2.3.1 Preparation of DNA of esiRNA Region

The pre-cloned full-length DUBs plasmids and the esiRNA region-containing plasmids essentially represent two different sources of esiRNA region. For the former, esiRNA regions were PCR-amplified from the plasmids with corresponding primers (Fig 5.6). Occasionally, troubleshooting needs to be done to optimize PCR conditions, including annealing temperature, primer concentration and plasmid concentration. PCR products were purified from gel after DNA agarose gel electrophoresis using Gel Extraction Kit from QIAGEN according to the manufacturer's instruction.

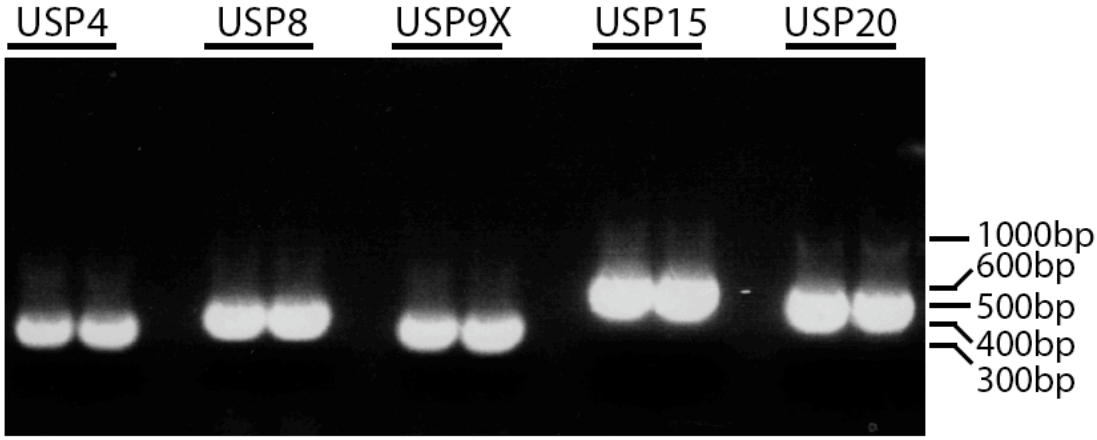

Figure 5.7. **PCR amplification of esiRNA region from precloned plasmids.** High fidelity DNA polymerase, Pfu was used for PCR reaction and reactions were set up in duplicates. Samples were analysed by DNA agarose gel electrophoresis.

On the other hand, the esiRNA-region containing plasmids were multiplied by growing the transformed bacteria in culture medium and the plasmids were subsequently purified using the QIAGEN Plasmid MiniPrep Kit. These plasmids cannot be used immediately for *in vitro* transcription because pCR4TOPO vectors contain a T7 promoter sequence 70bp downstream of the PCR insert, which is flanked by a T7 promoter sequence on both ends. This may result in a mixture of IVT product. Therefore, the esiRNA regions had to be isolated from the plasmids. This

was achieved by digestion using EcoRI restriction enzyme, which cleaves the plasmid backbones 5-base pairs upstream and downstream of the esiRNA regions. Following restriction digest, the reaction mixtures were separated by DNA agarose gel electrophoresis (Figure 5.7) and the desired esiRNA regions were purified using the Gel Extraction Kit from QIAGEN.

Concentrations of the purified esiRNA regions were measured using a wavelength NanoDrop 2000. During each round of such preparation, the DNA of esiRNA regions purified are enough for at least 10 *in vitro* transcription reactions and can be kept at -20˚C for long term storage.

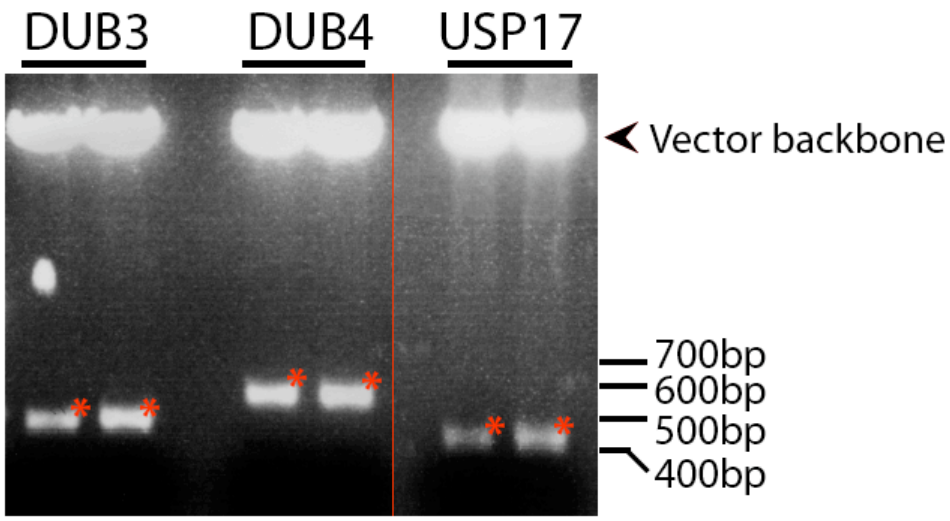

Figure 5.8. EcoRI digestion of pCR4TOPO plasmids containing esiRNA region. 3µg plasmids were incubated with EcoRI at 37˚C for 2 hours and samples were resolved on DNA agarose gel. Marked with asterisks are the esiRNA regions.

5.2.3.2 *In vitro* Transcription

To generate double stranded RNA (dsRNA) from the purified DNA of esiRNA region by *in vitro* transcription (IVT), the T7 MEGAScript® Kit from Ambion was used. Reactions were set up as was described by Kittler *et al.* in 2007 (see Chapter 2 for details). The IVT reaction involves binding of T7 RNA polymerase to the T7 promoter

sequence on the 5' end and subsequent transcription of the esiRNA region. The two resulting single-stranded RNAs, which are complementary to each other, are allowed to anneal to each other to give dsRNA. Then, 0.4µl of IVT product was analysed by DNA Agarose gel to check efficiency of the transcription reaction and the rest was used for RNAse III digestion.

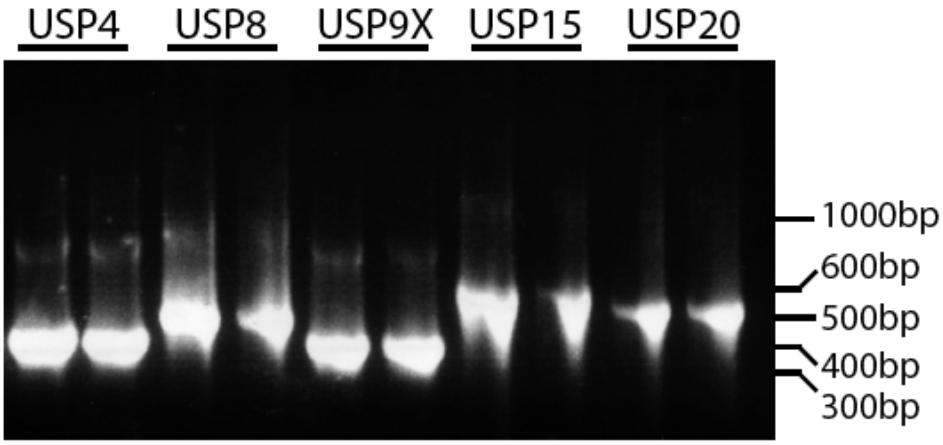

Figure 5.9. ***In vitro*** **transcription of esiRNA region.** IVT reaction was set up with 4µl of PCR product and 6µl of reaction mixture from MEGAScript® T7 Transcription Kit (Ambion). 0.3µl of reaction product was analysed on DNA agarose gel.

5.2.3.3 RNAse III Digestion

For enzymatic digestion of the long dsRNA resulting from IVT into short fragments of esiRNAs, the enzyme Ribonuclease III (RNAse III) from *Escherichia coli* was used. We have received a kind donation from Dr. Dun Yang (UCSF) for the plasmid of RNAse III with a glutathione S transferase (GST) epitope tag (pGEX2T-RNAse III). The sequence of the plasmid was verified and the plasmid was transformed into BL21 competent cells for protein expression. To purify GST-RNAse III, I have set up 1L overnight culture of BL21 cells and GST-RNAse III expression was induced by isopropyl-β-D-thio-galactoside (IPTG; 0,4mM) for 3 hours. After that, bacteria cells were pelleted by centrifugation and lysed by sonification. The cell lysate was then spun at 55000 rpm and GST-RNAse III, which is a soluble protein, was pulled down from the supernatant using glutathione-agarose beads. Next, GST-

RNAse III was retrieved sequentially from the beads and the eluates were combined and dialysed for further purification. Samples were taken at every step during this process and analysed by SDS-PAGE to check expression of GST-RNAse III, purity of the pulled down product and the extent of degradation. As shown on figure, GST-RNAse III was also found in the cells pellet while remaining largely soluble, and there was no significant degradation of the protein detected. Finally, the protein concentration of the purified GST-RNAse III was determined, diluted to 1.5µg/µl and stored at -80˚C in 5µl aliquots.

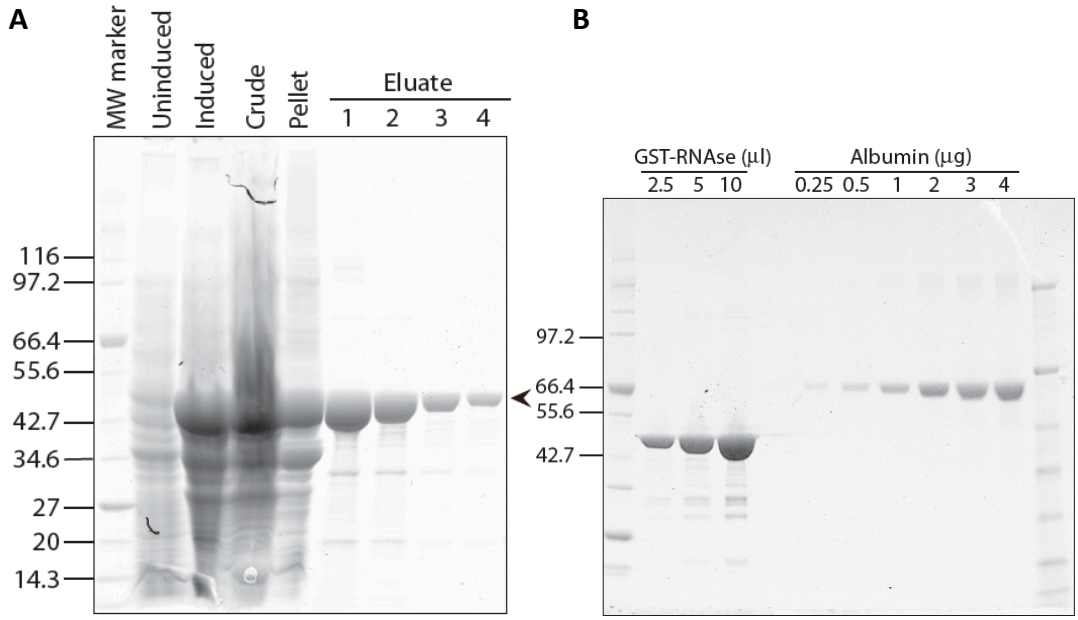

Figure 5.10. **Purification of GST-RNAse III. (A)** GST-RNAse III was induced in BL21 cells and was purified from supernatant of the subsequent bacterial lysates, by incubation with glutathione agarose beads at 4˚C for 3 hours. Bound GST-RNAse III was eluded sequentially with 1ml elution buffer. Samples were taken for bacterial cells before and after induction, crude lysates after sonification, pellet after centrifugation of lysate and the eluates, and were analysed by SDS-PAGE.(Note: Arrow = GST-RNAse III). (**B).** Eluates were combined and dialysed against 4L dialysis buffer. Dialysed samples were analysed by SDS-PAGE with different loading volume alongside increasing amount of albumin to determine amount of GST-RNAse III.

The dsRNA generated by IVT was incubated with purified GST-RNAse III in an esiRNA digestion buffer, with a total volume of 100µl. The digestion mixture was kept shaking at 900 rpm, and for 4 hours at 25˚C followed by 2 hours at 37˚C. 3µl of

the digested samples was analysed by DNA agarose gel electrophoresis to ensure complete digestion, with a single band running around 25bp.

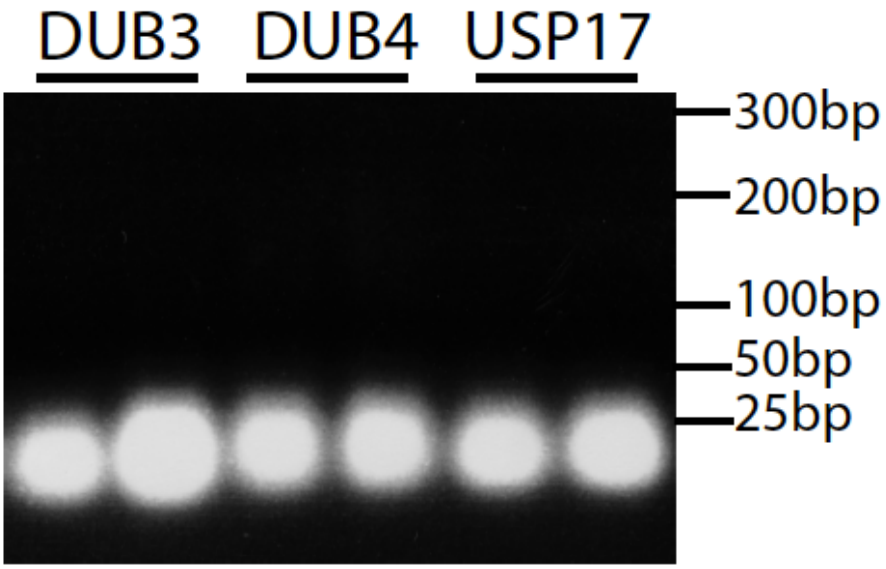

Figure 5.11. **RNAse III digested dsRNA.** IVT products were incubated with 6µg of RNAse III and in esiRNA digestion buffer. Samples were kept shaking at 900rpm for 4 hours at 25˚C followed by 2 hours at 37˚C. 3µl of digested samples were resolved on 2% DNA agarose gel at 100V for 1 hour.

5.2.3.4 esiRNA Purification

The digested esiRNAs were purified by ion exchange column and precipitated with isopropanol. The esiRNA was spun down by centrifugation and dried in a SpeedVac. The resulting pellet was resuspended in 100µl of RNAse-free water and concentration was measured using NanoDrop. 50µl of each of the esiRNAs were distributed on 2 96-well plates (Table 5.4), and these serve as a resource for future esiRNA DUB library screens.

Table 5.3

	1	2	3	4	5	6	7	8	9	10	11	12
A	CYLD	DUB3	DUB4	USP1	USP2A	USP3	USP4	USP5	USP6	USP7	USP8	USP9X
B	-	USP10	USP11	USP12	USP13	USP14	USP15	USP16	USP17	USP18	USP19	USP20
C	USP21	USP22	USP24	USP25	USP26	USP27X	USP28	USP29	USP30	USP31	USP32	USP33
D	USP34	USP35	USP36	USP37	USP38	USP39	USP40	-	USP42	USP43	USP44	USP45
E	USP46	USP47	USP48	USP49	USP50	USP51	USP52	USP53	USP54	USPL1	BAP1	UCHL1
F	UCHL3	UCHL5	ATXN3	ATXN3L	JOSD1	JOSD2	BRCC3	CSN5	CSN6	EIF3H	EIF3S5	MPND
G	MYSM1	PRPF8	PSMD 14	PSMD7	STAM BP	STAM BPL1	OTUB1	OTUB2	OTUD1	OTUD3	OTUD4	OTUD5
H	OTUD6 A	OTUD6 B	OTUD7 A	OTUD7 B	PARP11	TNFAIP 3	VCPIP1	YOD1	ZRANB 1	GFP		

Table 5.3 esiRNA plate map.
Note : Generation of esiRNA for USP9X and USP41 was unsuccessful; esiRNA for GFP was generated as a non-targeting control.

5.2.4 Characterisation of theDUB esiRNA library

Since the DUB esiRNA library is a newly introduced resource in the lab, characterisation works were carried out to determine the knockdown efficiency of chosen esiRNAs, consistency of the esiRNAs, optimisation of the knockdown protocol and to compare knockdown efficiency across a variety of cell lines.

5.2.4.1 Checking knockdown efficiency of esiRNA

The first obvious issue to address is the knockdown efficiency of individual esiRNA. With the limited resource (in term of antibody availability and cost) and time, it was not feasible to determine the knockdown efficiency of esiRNA generated for each and every DUB. I have therefore selected a panel of esiRNAs for which the knockdown efficiency was to be assessed, either by immunoblotting (IB) to check

protein level (if the corresponding antibodies are available in the lab) or RT-PCR to check mRNA level.

I first assessed the silencing efficiency of several DUBs in MCF7 cells by IB (Figure 5.12). Transfection was performed with 300ng of esiRNA per 1ml of medium, which is equivalent to 40nM of total oligonucleotide. Cells were lysed 72 hours post-transfection and equal amount of lysates were loaded for SDS-PAGE, followed by IB. Densitometry analysis was then performed using Image J software to determine relative amounts of DUB protein level between control and knockdown samples. All measurements were normalised to tubulin level.

In all cases, there was a visible depletion of individual DUBs following esiRNA transfection. Quantification revealed varying knockdown efficiency: USP8 – 75%, USP15 – 70%, USP20 – 70%, USP4 - ~65%, USP47 - ~65%, USP9X - ~60%, USP38 – ~50%, with more than 50% of proteins depleted in general.

In parallel to this, RT-PCR experiments were carried out to determine silencing efficiency of several other DUB esiRNAs. For this purpose, I have designed specific primers for a panel of DUBs summarised in Table 5.5. Primer pairs were designed to amplify regions not greater than 200bp and to span introns wherever possible to avoid amplifying genomic DNA, which may be present as contaminants in mRNA samples.

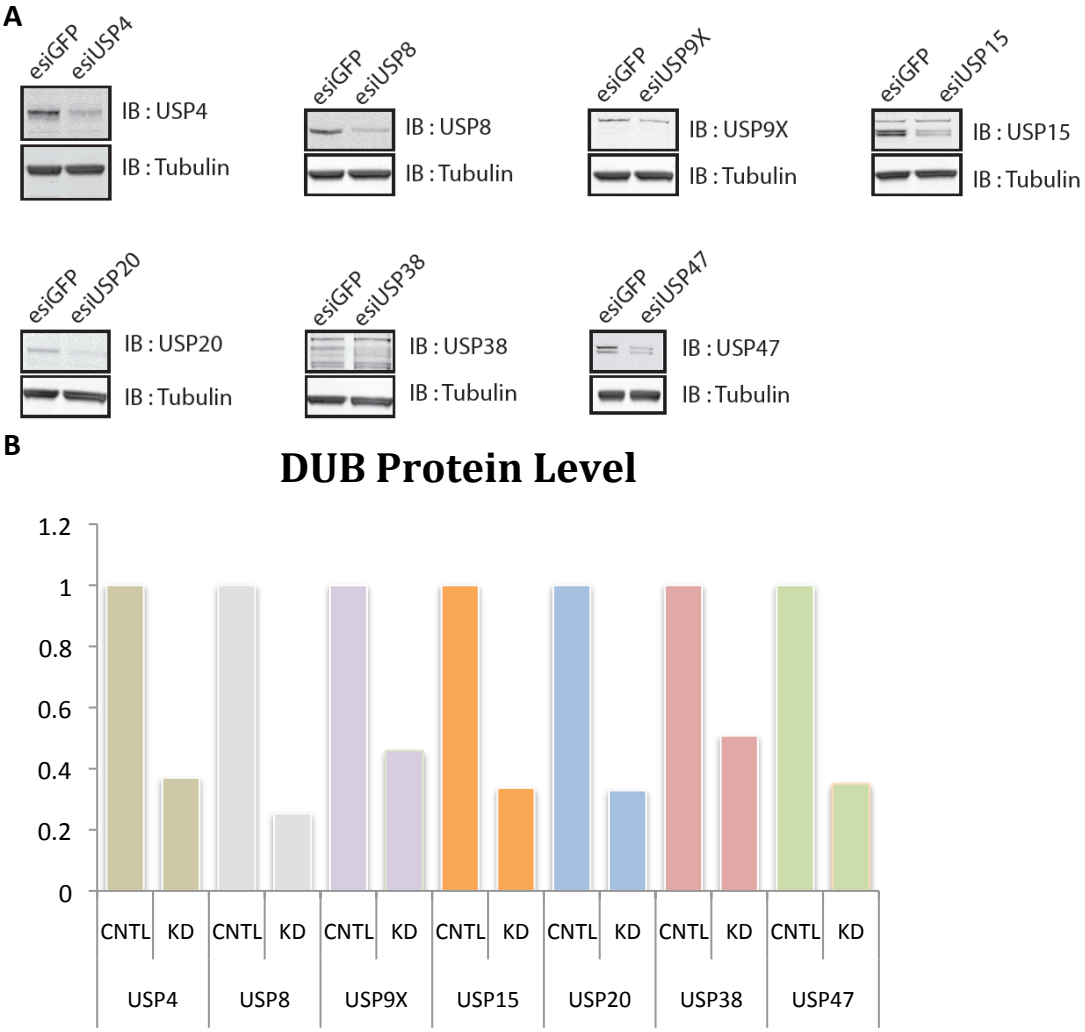

Figure 5.12. **Silencing efficiency of selected esiRNAs determined by IB. (A)**MCF7 cells seeded in 6-well plates were transfected with 100ng (in 1ml DMEM medium) of esiRNA targeting individual DUBs or esiRNA against GFP as negative control. Cells were lysed 72 hours post-transfection, resolved by SDS-PAGE and transferred onto nitrocellulose membrane for IB with individual DUB antibodies and tubulin. Images were acquired by infra-red scanner (Odyssey, LICOR). **(B)** Densitometric analysis was performed using Image J software to determine band intensity and quantification was done by normalisation to tubulin and to negative control.

Table 5.4.

DUBs	Primers	Sequence	Exon	Amplicon size / bp
USP21*	Fwd	ACACGTGCTTCCTGAATGCT	-	128
	Rev	AGGCTTCAGTGAGCTCTTGG	-	
USP27X	Fwd	CATTTAGCAGGGTACAGGCAA	1	128
	Rev	GATGCAGTTACAGTGGTTGGG	1	
UCHL3	Fwd	GGGACAAGATGTTACATCATCAG	4	126
	Rev	GGTTGATCCAGATTCAAAGTGC	4-5	
ATXN3	Fwd	GTATGCAAGGTAGTTCCAGAAAC	7-8	126
	Rev	TGTTGCTGCTTTTGCTGCTGT	8	
JOSD2	Fwd	GAACCCTCATCGCAGCCTC	2-3	148
	Rev	GATCAGCCCCAGTACCTGG	3-4	
CSN5	Fwd	CTGGACTAAGGATCACCATTAC	1-2	107
	Rev	GACCCATCACTTCCAAGTTGC	2	
MPND	Fwd	GTTCTGGCAGGGGTCTCAG	3-4	121
	Rev	GAACCCGGATCTTGCTGTC	4	
PRPF8*	Fwd	TATGACCACCAGCCGTTGAG	-	193
	Rev	CCATATTGAGTGCCTTGGACG	-	
PARP11	Fwd	GCGAATCCAGAGATGTTTCAC	1-2	148
	Rev	GTTGGTATCCGGCTGAAACATG	2-3	
STAMBPL1*	Fwd	AACCTTATTTGTAGAAAAGCTTCCTAA	-	101
	Rev	GGAATGCAATCTCCTTCAGTTTC	-	
OTUD1*	Fwd	CCGACCATCTCGACCACTTC	-	251
	Rev	GTGTCCGTTACTGAGCCAACTG	-	
TNFAIP3*	Fwd	TGCCCAGGAATGCTACAGATAC	-	228
	Rev	TGGACGGGGATTTCTATCACC	-	

Table 5.4 DUB RT-PCR primers.
*Primers already available in lab.

In this set of experiments, mRNA of MCF7 cells was harvested 72 hours post-transfection of esiRNA and was reverse-transcribed to give cDNA for RT-PCR reactions. Varying silencing efficiency was observed across the panel of DUBs: more than 90% for OTUD1, MPND and JOSD2; around 80% for TNFAIP3, PARP11 and STAMBPL1; 50-60% for USP21, PRPF8 and ATXN3.

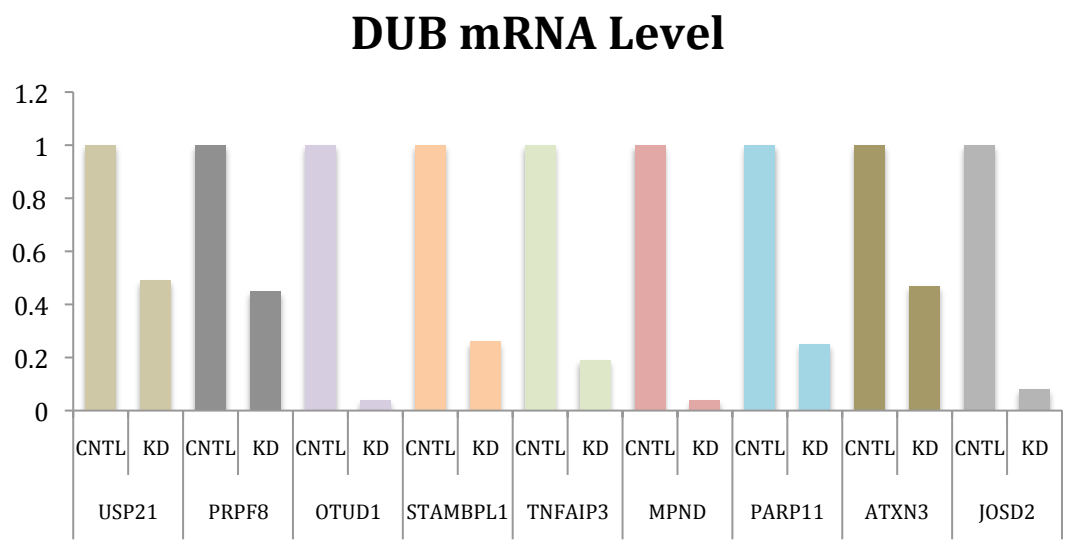

Figure 5.13. **Silencing efficiency of esiRNA determined by RT-PCR.** mRNA was extracted from MCF7 cells 72 hours post-transfection using RNA extraction kit (RNeasy Mini Kit, QIAGEN) and 1µg of mRNA was reverse-transcribed to give cDNA. RT-PCR reactions were set up in triplicates, with β–actin as a reference gene and SYBR green reaction mixture (BioRad) was used. RT-PCR results were acquired using an iQ5 Real-Time PCR system (BioRad). mRNA levels of DUBs were normalized to that of the control sample.

Among the esiRNAs for which the silencing efficiency was assessed, loss of DUBs at protein level or mRNA level was observed, demonstrating the ability of esiRNA to knockdown gene of interest. However, in most cases, silencing efficiency was not reaching the desirable efficiency of more than 80%. This can possibly be explained by the fact that esiRNA comprises of a complex mixture of siRNAs of heterogenous silencing potency due to varying sequence and lengths. Such variation in silencing capacity is not uncommon among chemically synthesized siRNAs. In light of that, higher amount of esiRNA might have to be used to achieve optimal silencing efficiency. Besides, in cases where silencing efficiency was determined using IB, longer half-lives of proteins of interest may mask the effect of knockdown of the corresponding genes. Therefore, proteins have to be harvested from cells after a longer time period post-transfection for the knockdown effect to be more pronounced as assessed by IB.

5.2.4.2 Titration of amount of esiRNA

To optimize the amount of esiRNA used for knockdown, I have performed knockdown of USP8 and USP38 in MCF7 cells with varying amount of esiRNA, from 25ng to 300ng per ml of medium at the scale of a 6-well plate (Figure 5.14).For both DUBs, there was a dose-dependent effect of esiRNA, where increasing amount of esiRNA resulted in better silencing effect. It was quite remarkable that with only one-twelth (25ng) of the recommended amount (300ng) of esiRNA, a silencing efficiency of 50-60% was achieved for both DUBs. Knockdown with higher amount of esiRNA (>400ng/ml) was not possible with MCF7 cells because substantial cell death was observed, demonstrating the potential toxicity of esiRNA.

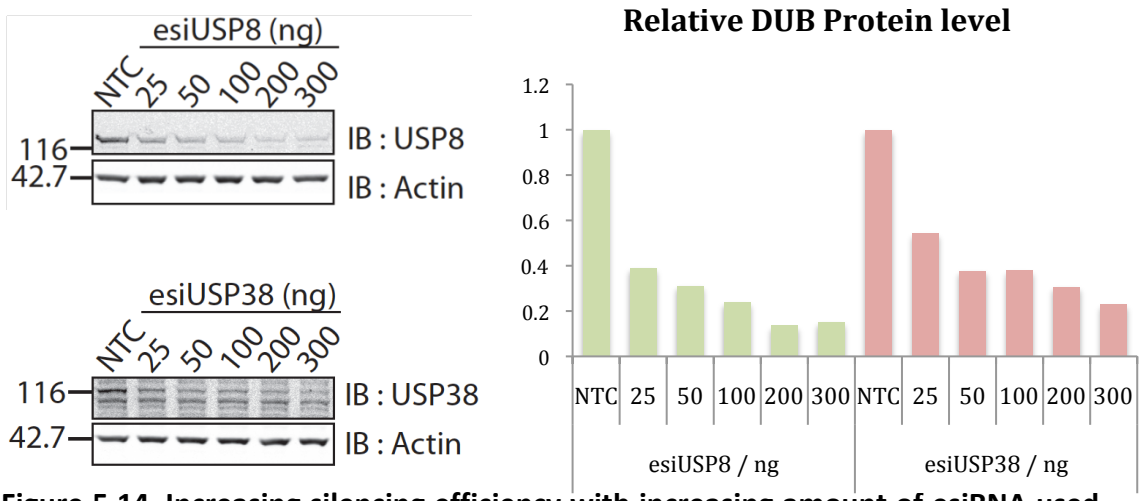

Figure 5.14. Increasing silencing efficiency with increasing amount of esiRNA used. MCF7 cells were transfected with varying amount of esiRNA targeting USP8 or USP38. Proteins were harvested from cells 72 hours post-transfection and analysed by SDS-PAGE, followed by IB. Densitometric analysis was performed using Image J software to determine relative amount of protein and amount of DUB present was normalized to actin and to the control.

I have also titrated PRPF8 esiRNA required for optimal knockdown in MCF7 and A549 cells, by checking the mRNA level (Figure 5.15). Although a dose-dependent effect of PRPF8 esiRNA was observed in both cell lines, PRPF8 silencing was more efficient in MCF7 than in A549. Increasing esiRNA amount beyond 300ng did not result in better silencing efficiency.

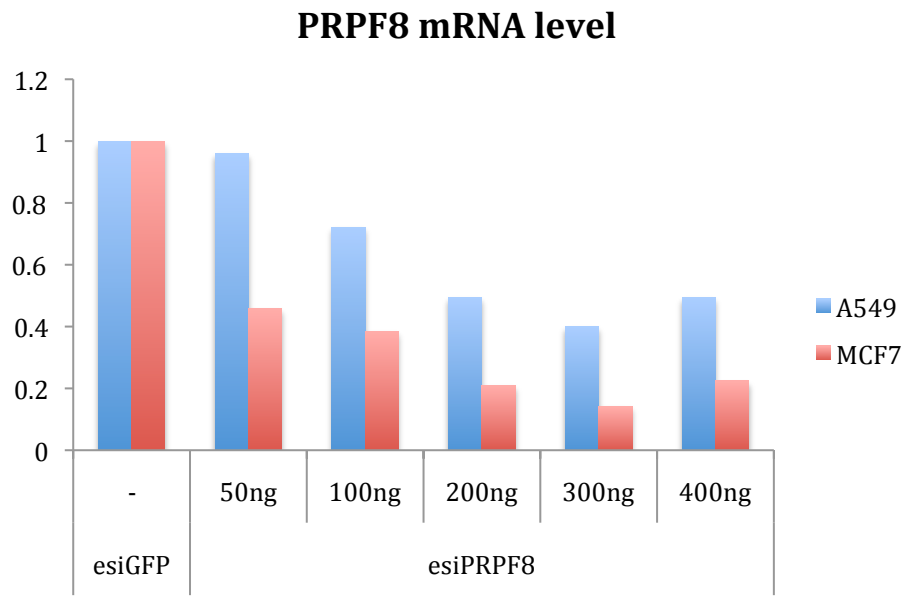

Figure 5.15. **Increasing silencing efficiency with increasing amount of esiRNA used.** MCF7 cells were transfected with varying amount of PRPF8 esiRNA. Total mRNA was harvested from cells 72 hours post-transfection and mRNA level of PRPF8 was assessed by QPCR.

Once again, these experiments highlighted the varying silencing capacity of esiRNA observable at different amount of esiRNA used. Titration of amount of esiRNA for individual DUBs is therefore recommended to achieve optimal knockdown. However, in the context of a screen, this is not feasible and the maximal dose of esiRNA a particular cell line can tolerate should serve as a guideline for amount of esiRNA used.

5.2.4.5 Comparison of esiRNA knockdown efficiency across cell lines

As shown in Figure 5.15, knockdown efficiency of esiRNA varies between cell lines. To assess this in a more comprehensive manner, I have performed knockdown of USP8 using esiRNA in another 3 cancer cell lines, namely A549, SW480 and HeLa (Figure 5.16). In all cases, a knockdown efficiency of greater than 75% was achieved, with the highest knockdown efficiency observed in HeLa cells.

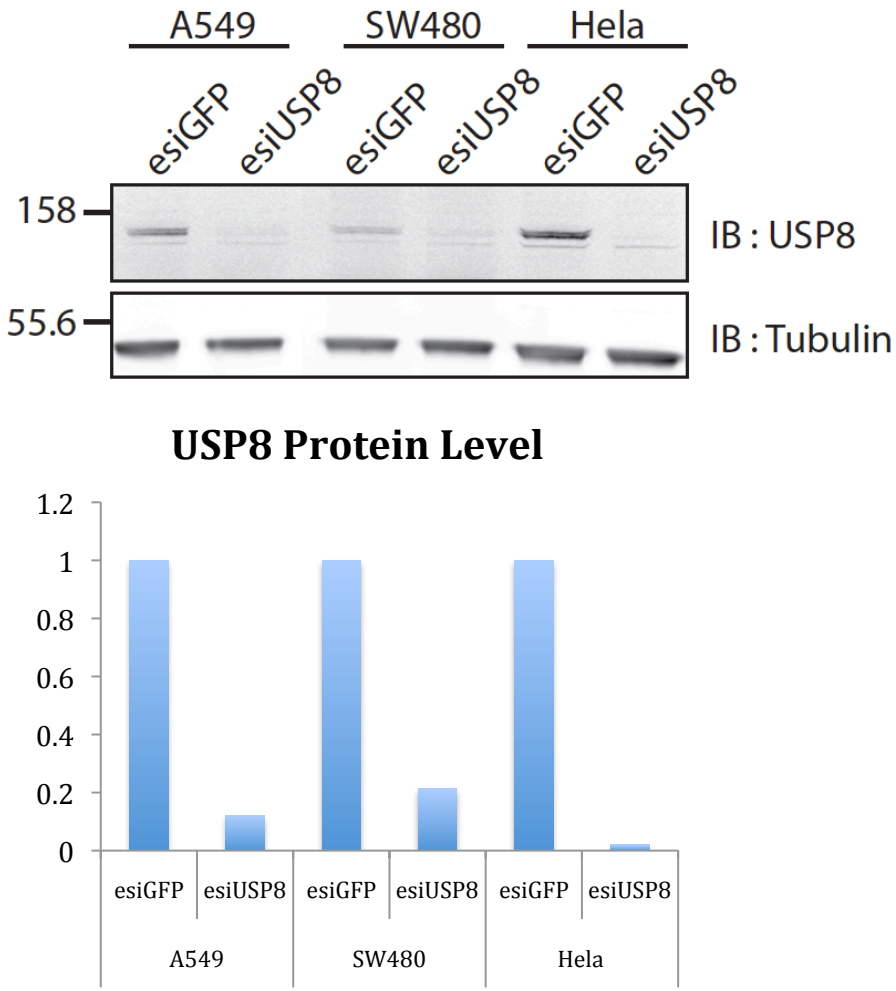

Figure 5.16. Consistent knockdown efficiency of USP8 across different cell lines. Cells were transfected with 300ng USP8 esiRNA. Proteins were harvested from cells 72 hours post-transfection and analysed by SDS-PAGE, followed by IB. Densitometric analysis was performed using Image J software to determine relative amount of protein and amount of DUB present was normalized to actin and to the control.

5.2.4.6 Determining consistency of esiRNA

Since the digestion of dsRNA into esiRNA by RNAse III is a random process, there is hence a concern of consistency of esiRNA prepared from different batches. To address this, I prepared several batches of esiRNA for several DUBs with the help of a summer student and assessed the knockdown efficiency in SW480, a colon cancer cell line. No obvious difference in silencing efficiency was observed between esiRNAs prepared from different batches (Figure 5.17). In other words, despite the

the fact that dsRNA digestion is random, the consistency of the esiRNAs generated is preserved across batches.

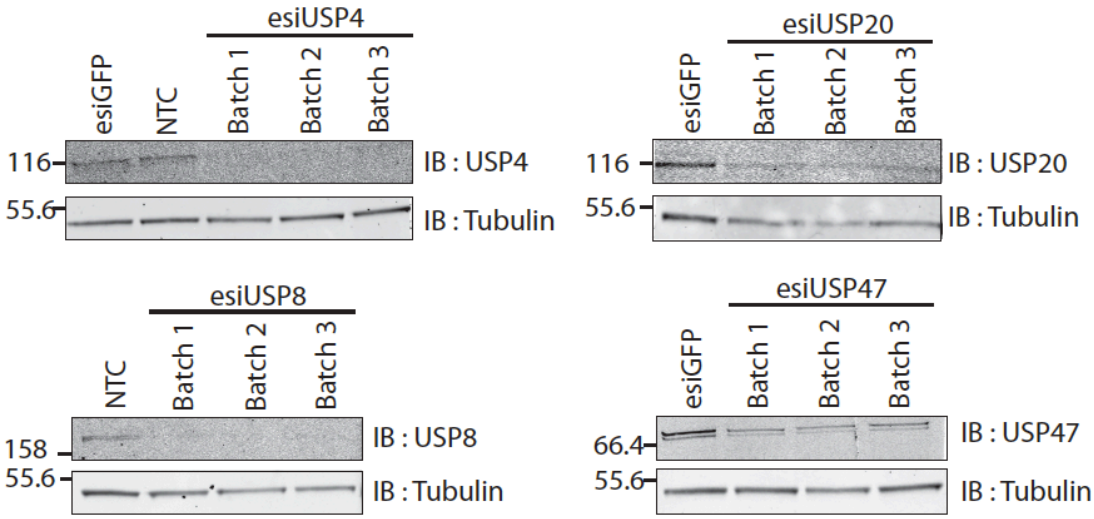

Figure 5.17. **Different batches of esiRNA lead to consistent knockdown efficiency.** SW480 cells were transfected with 3 different batches of esiRNA at 300ng/ml for each set of knockdown and cells were lysed 72 hours post-transfection. An additional negative control using QIAGEN non-targeting siRNA (NTC) was also set up. Lysates were resolved by SDS-PAGE and transferred to nitrocellulose membrane followed by IB of DUBs. Tubulin was used as loading control. Images were acquired by Infra red scanner (Odyssey, LICOR).

5.2.5 Validation of known DUB-protein functional relationship

As mentioned before, one of the aims to generate a DUB esiRNA library is to provide an alternative tool to validate known DUB-protein functional relationship determined by siRNA knockdown. For this purpose, esiRNA against BAP1 was used for knockdown in MCF7 to reconfirm functional relationship between BAP1 and β-catenin as was identified in Chapter 4. Similar to knockdown using BAP1 siRNA, depletion of BAP1 using esiRNA also resulted in loss of β-catenin at the protein level and the knockdown efficiency was as good as knockdown using BAP1 siRNA (Figure 5.18).

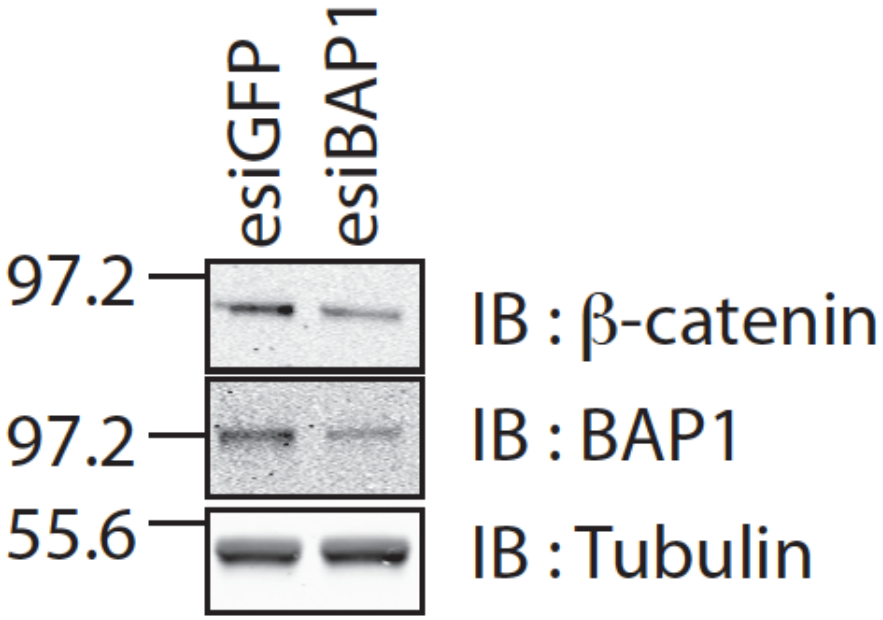

Figure 5.18. Knockdown of BAP1 using esiRNA resulted in β-catenin depletion. MCF7 cells were transfected with esiRNA against GFP or BAP1 at 300ng/ml and were lysed 72 hours post-transfection. Lysates were resolved by SDS-PAGE and transferred to nitrocellulose membrane followed by IB of DUBs. Tubulin was used as loading control. Images were acquired by Infra red scanner (Odyssey, LICOR).

5.3 Discussion

The discovery of RNA interference (RNAi) by Andrew Fire and Craig Mello in 1998 (Fire et al., 1998) has sparked immense interest and intensive investigations to understand this well-conserved cellular mechanism. The knowledge derived from these works is exploited extensively in various technological platforms, ranging from functional genomics to biotechnology to medicine.

RNAi is now a routine experimental technique employed in the attempts to annotate gene functions. siRNA library screens are regularly performed to identify genes involved in particular cellular processes in an unbiased fashion. As powerful as the technology is, and despite the state-of-the-art algorithms developed to predict specific siRNA, there is a persistent problem of off-target effects, where a non-desired gene is silenced. To address this issue, pools of siRNAs comprising of 2-4 siRNAs are normally used, but off-target effects due to 1 siRNA, which dominates in

the pool, is sufficient to mask the effects of others. In the case of having a false positive hit, deconvolution experiments, which involve knockdown using individual siRNAs comprise the pool, can help reveal the off-target effects. However, especially in the context of a screen, a true target will be missed if the off-target effects result in a false negative output. Not to mention these validation experiments are cost- and time-intensive.

In an effort to develop an alternative strategy, which can potentially circumvent the above-mentioned problems and is available at relatively economic price, Frank Buccholz and colleagues have developed esiRNA. In essence, esiRNA is a very complex pool of siRNAs generated at random following enzymatic digestion of dsRNA of a carefully selected region within the transcript of a particular gene. It is the complexity of this pool that is claimed to minimize off-target effects, and hence gives greater specificity. To demonstrate this, Kittler et al. silenced MAPK14 using chemically synthesized siRNA, either individually or with an increasing number of siRNAs within a pool of 2 to 12, where concentration of individual siRNAs are diluted with increasing complexity of the pool,and performed a genome-wide microarray expression profliling (Kittler *et al.*, 2007). A less wide spread effect on gene expression profile was observed with increasing complexity of pool, indicating a much lower extent of off-target effects. This is because in a very complex pool, individual siRNA is present at such a low level that, if any, off target it causes is trivial.

Given the pitfalls of siRNAs and potential promises of esiRNAs, we have decided to adopt this technology, as a complement to the siRNA technique and eventually as a resource to the research community. Compared to the genome wide esiRNA library generated (Kittler *et al.*, 2007), the construction of a esiRNA DUB library consisting of 93 esiRNAs represents a more humble effort.

This chapter provides a detailed description of the works that have been carried out to achieve this aim. The most important and the very first step was to design specific primers against desired esiRNA region. Firstly, I extracted information

of predicted optimal esiRNA region for individual DUBs from the online database RiDDLE. Then, I checked whether it was possible to amplify these esiRNA regions from the ORFs already cloned in the lab by sequence analysis. Next, I checked the cross-silencing activity of the esiRNA regions by doing a BLAST analysis. Although DEQOR was designed to penalize esiRNA regions with high cross-silencing activity, unexpectedly, a few of these esiRNA regions showed extensive overlapping region with the other DUBs. In the most extreme case, the predicted esiRNA region for USP6 was completely identical to the coding sequence of its close relative, USP32. This highlighted potential limitations with the DEQOR software and/or the validity of information available on RiDDLE. Therefore, in the construction of esiRNA against a family of proteins such as DUBs, it is ideal to manually check the cross-silencing activity of individual esiRNA regions. Having mentioned that, it is however quite challenging and sometimes impossible to generate specific esiRNA against proteins, which are closely related, such is the case of USP9X and USP9Y. For esiRNA regions predicted to have low efficiency, I tried to select alternative regions, which are more efficient. Although it was only possible for some of the DUBs, this once again raised the question of reliability of information available on the RiDDLE database. Finally, I have optimized the primer sequence to ensure optimal qualities of primers for PCR amplifications. The finalized esiRNA regions represented the best compromise between esiRNA efficiency, specificity, primer qualities and coverage that was possible. Following the designs of primers, the generation of esiRNA was relatively straight forward, although trouble-shootings were required occasionally, especially at the stage of PCR amplifications. At the end of the process, I have only managed to produce esiRNAs for 91 out of 93 DUBs. Generation of esiRNA for USP9Y and USP41 was challenging, because amplification of esiRNA region from cDNA was difficult. It was also possible that the genes were not expressed in MCF7 and A549 cells, from which the cDNAs were prepared.

Following generation of the esiRNAs, I have selected some of them, based on the resource available in the lab, and tested their silencing efficiency and consistency. All the esiRNAs tested were able to induce silencing of their targets, but with varying efficiency, from 50 – 95%. The unsatisfactory silencing efficiency for

some of the esiRNAs can be a concern as the effect of knockdown might not be prominent with residual amount of enzymes left in the cells. The silencing efficiency can possibly be optimized by increasing amount of esiRNA or extending the length of experiment. However, there are potential issues associated with toxicity of higher amount of esiRNA and longer time of transfection, and this is also dependent on the type of cell lines used. Moreover, as mentioned before, optimization of individual DUB esiRNAs is not possible in the context of an esiRNA screen.

Following the characterization of the esiRNA DUB library, this resource can be included into the pipeline of current research projects in the lab, either for esiRNA library screen or as an additional alternative for validation of hits identified from siRNA library screen. Besides, in the case of large-scale studies such as proteomics, esiRNA is a much cheaper alternative. We anticipate with a potentially higher specificity conferred by esiRNA, the annotation of functions of DUBs will become more efficient.

Chapter 6

Characterisation of USP38

6.1 Introduction

USP38 was identified as a potential DUB involved in regulation of E-cadherin. USP38 is 1 of 56 members of the USP family and is 1042 amino acids in length. Other than the catalytic domain, there is no other predicted domain (Figure 6.1). The biological function of USP38 is still yet to be annotated. One way of assigning biological function of a gene is to identify interacting partners of the gene product. By immunoprecipitation of epitope-tagged DUBs expressed in stable cell line followed by mass spectrometry analysis, Sowa and colleagues attempted to identify the interactome of 75 human DUBs (Sowa *et al.*, 2009). This report identified 5 proteins, which co-immunoprecipitated with USP38, namely 60S ribosomal protein L7 (RPL7), 40S ribosomal protein S12 (RPS12), Galectin 7 (LGALS7), heat shock 27kDa protein 1 (HSPB1) and homeo box (H6 family) 3 (HMX3). Table 6.1 summarises the known functions of these proteins. While their functions may be regulated by USP38, none of these potential interactions has been followed up and further verified.

Figure 6.1 Domain architecture of USP38. USP38 is 1042 amino acid in length and amino acids 442-946 correspond to the catalytic USP domain of the protein.

As discussed in Chapter 3, I did not manage to decipher the mechanism by which USP38 can affect E-cadherin regulation. Since the biological function of USP38 is unknown, I was very interested in making a contribution towards the knowledge

about this protein and I anticipated revealing aspects of its functions which may be relevant to E-cadherin regulation.

The works presented in this chapter aim to characterise USP38 by (i) sequence analysis and (ii) creating truncation mutants to see effect on USP38 localisation.

Table 6.1

Gene name	Description of Function	Reference
RPL7	Component of 60S ribosome; involved in regulation of translation machineries.	Seshadri *et al.*, 1993.
RPS12	Component of 40S ribosome.	Herault *et al.*, 1991.
LGALS7	Carbohydrate binding protein implicated in various cellular processes including cell adhesion, cell-matrix interaction, apoptosis and differentiation.	St-Pierre *et al.*, 2012.
HSPB1	Molecular chaperones that are overexpressed in stress conditions.	Arrigo *et al.*, 2007
HMX3	A transcription factor essential for development of inner ear and successful pregnancy.	Wang *et al.*, 1998.

Table 6.1 Function of proteins co-immunoprecipitated with USP38. The USP38 interacting partners were identified by Sowa and colleagues using mass spectrometry following immunoprecipitation of HA-tagged USP38 in HEK293T stable cell line expressing HA-USP38 (Sowa *et al.*, 2009). The references listed represent literature reports of the function of each protein.

6.2 Results

6.2.1 Sequence Analysis of USP38

USP38 is found expressed in a number of higher eukaryotes, including human (Homo sapiens), cow (Bos taurus), chicken (Gallus gallus), rat (Rattus norvegicus) and mouse (Mus musculus). By performing a NCBI conserved domain search, an insertion of 46 aa, corresponding to residues 538 – 583 in the catalytic domain, unique to the human USP38 was found (Figure 4.2).

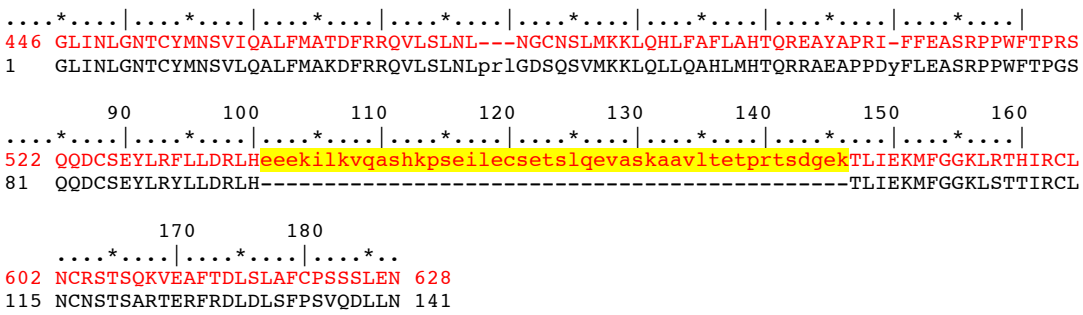

Figure 6.2 USP38 USP domain contains an insertion of 46 amino acids. A NCBI conserved domain search was performed for the USP domain of USP38. Amino acid residues in red belong to human USP38 and amino acid residues in black are residues conserved among the USPs. The 46 amino acid residues highlighted in yellow represent an insertion unique to USP38.

By multiple sequence alignment of the catalytic domains of all the USPs, Ye and colleagues identified 6 boxes of amino acids residues, which are conserved across the USPs (Figure 4.3) (Ye *et al.*, 2009). USP38 has all the core components of the USP domain. It has a complete catalytic triad, comprising of the catalytic cysteine residue in Box 1, catalytic histidine residue in Box 5 and the catalytic asparagine residue in Box 6. Also, it has 2 Cys-x-x-Cys motifs for Zinc binding, in Box 3 and Box 4 respectively. The residues between the boxes of conserved amino acid residues within the catalytic USP domain represent inserted sequences. There is no amino acid insertion between Box 1 and Box 2, and the largest insertion of 90 amino acids is found between Box 3 and Box 4. Sequence alignment of USP38 of different species revealed that the region outside of the catalytic domains are also very well conserved across all the species and the insertions represent the only variable regions along the length of USP38 (see Appendix).

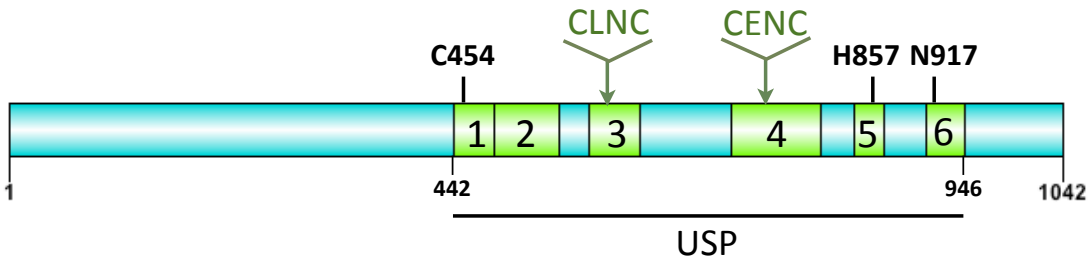

Figure 6.3 6 conserved boxes of amino acids across all human USPs in USP38. The numbered boxes indicate each of the 6 conserved boxes. The conserved boxes are interspersed with insertions of different sizes and the catalytic triad (C454, H857, N917) is indicated. The Cys-x-x-Cys motif for Zinc binding in box 3 and 4 are also depicted.

6.2.2 USP38 localises to the cytoplasm and nucleus.

I have generated several mutants of USP38, including a catalytically inactive mutant USP38 (USP38-C454A) and 5 truncation mutants, namely USP38-1-441, USP38-441-946, USP38-1-946, USP38-441-1042 and USP38-Δ538-583 (Figure 6.4). The USP38-Δ538-583 mutant has its insertion between box 2 and 3 of the USP deleted. Mutants which harbour deletion of other insertions were not created. The ORFs of these truncations were all cloned into GFP-expression vector and transfected into MCF7 cells.

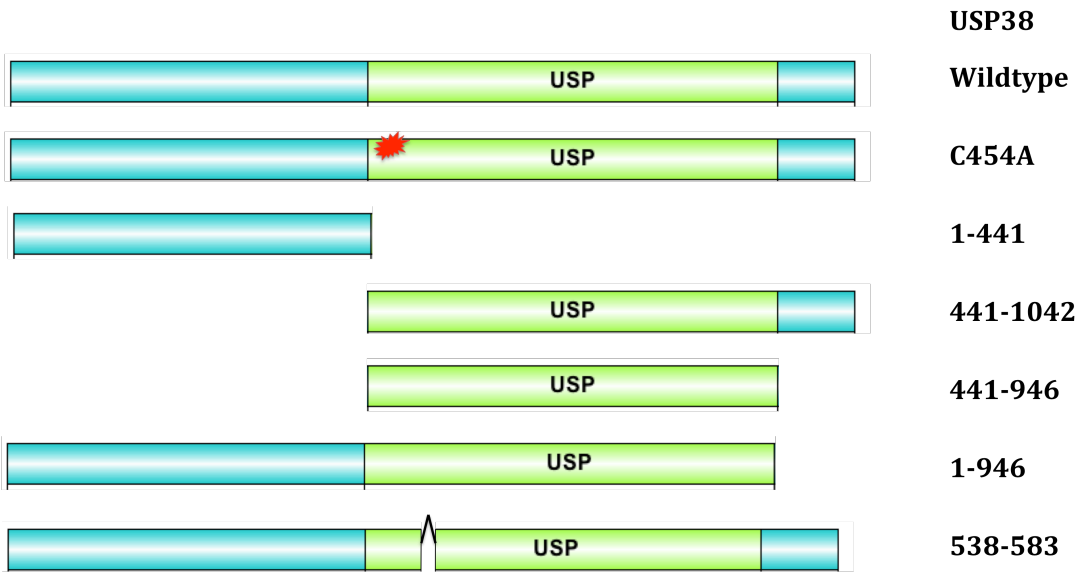

Figure 6.4 Truncation mutants of USP38.

Figure 6.5 shows the expression level of the different GFP-tagged USP38 proteins in MCF7 cells. The wildtype and catalytically inactive mutant of USP38 were expressed at similar level, with the mutant running at a slightly lower molecular weight. The USP38-1-441 and USP38-441-1042 mutants were expressed at higher level than the full-length proteins, while the expression levels of USP38-1-946, USP38-441-946 and USP38-Δ538-583 mutants were lower than the full length proteins. The difference in expression levels of these proteins could indicate differential stability of the proteins.

Since the stability of a protein can be regulated by post-translational modification such as ubiquitylation, I decided to check the ubiquitylation status of the different USP38 proteins (Figure 6.6). The wildtype USP38 showed low level of polyubiquitylation, whereas the C454A mutant was highly polyubiquitylated as indicated by the more intense higher molecular weight smear. The USP38-1-441 also showed high level of polyubiquitylation. Although its molecular weight is much smaller, the observed polyubiquitylation smear ran at the similar range to the USP38-1-946, which is about 500 amino acids longer, indicating that this truncation mutant was more heavily polyubiquitylated. The other truncations mutants, namely USP38-1-946, USP38-441-946 and USP38-Δ538-583, were also modified by

polyubiquitylation, among which the latter was most heavily polyubiquitylated. The truncation mutant USP38-441-1042 was the only protein that showed no detectable polyubiquitylation.

I then checked the localisation of the different USP38 proteins. Wildtype USP38 is found diffused in the cytoplasm and in the nucleus, but was excluded from nucleoli (Figure 6.7). The C454A mutant showed similar localisation in MCF7 cells. The truncation mutants USP38-1-441 was found expressed in the cytoplasm and the nucleus, but was not completely excluded from the nucleoli. The USP38-441-1042 mutant also showed similar localisation as the wildtype USP38. All the other truncation mutants, namely USP38-1-946, USP38-441-946 and USP38-Δ538-583 were found expressed only in the cytoplasm and were all excluded from the nucleus.

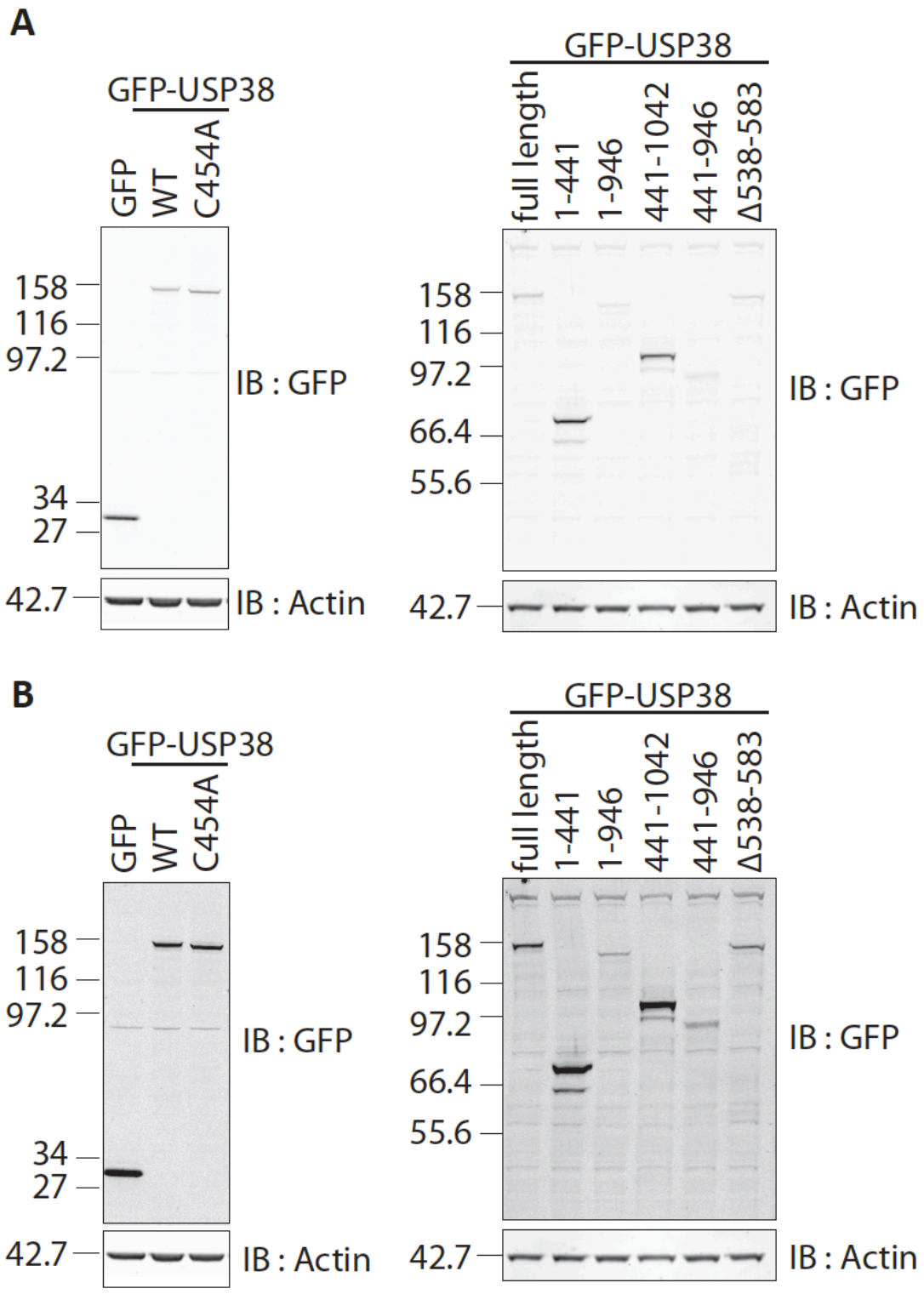

Figure 6.5 Expression level of full length and truncated USP38. MCF7 cells were transfected with the indicated plasmids and were lysed on the next day with NP40 lysis buffer. Lysates were resolved on 4-12% SDS-PAGE followed by transfer to a nitrocellulose membrane. The resulting blot was immunoblotted for GFP and actin. Blot images were acquired using an Odyssey laser scanner. Images in (A) and (B)

were from the same blot, and intensity of images in (B) was adjusted to allow visualisation of low expression level protein.

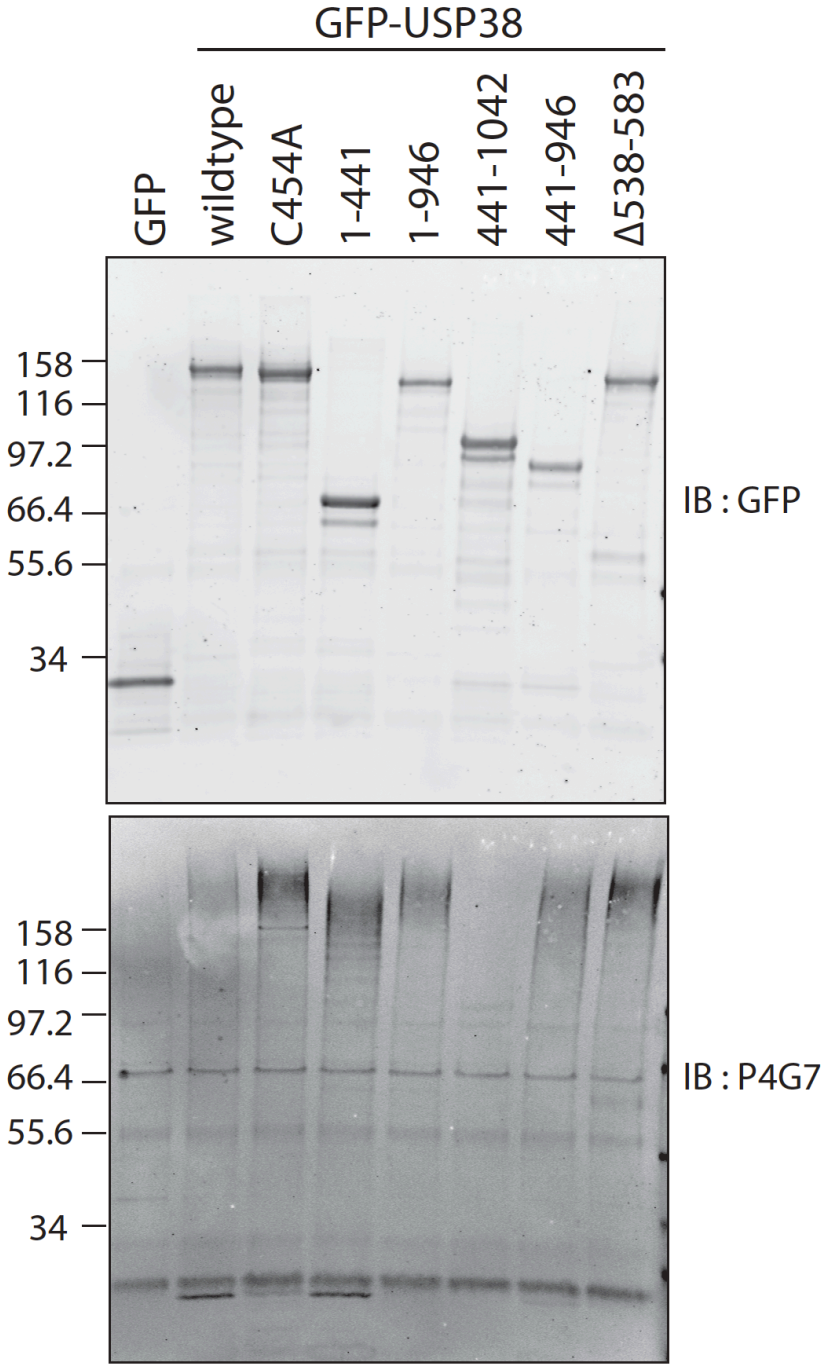

Figure 6.6 Ubiquitylation status of wildtype and mutant USP38. MCF7 cells were transfected with the indicated plasmids and were lysed on the next day with NP40 lysis buffer. For immunoprecipitation, 600µg of lysate was incubated with 1µl of GFP antibody and protein G sepharose beads in a final volume of 480µl in a 1.5ml Eppendorf tube, for 1.5 hours at 4˚C on a rotating wheel. Purified protein complexes were resolved by 4-12% SDS-PAGE followed by transfer to a nitrocellulose membrane. The resulting blot was immunoblotted for GFP and ubiquitin using P4G7 antibody. Blot images were acquired using an Odyssey laser scanner.

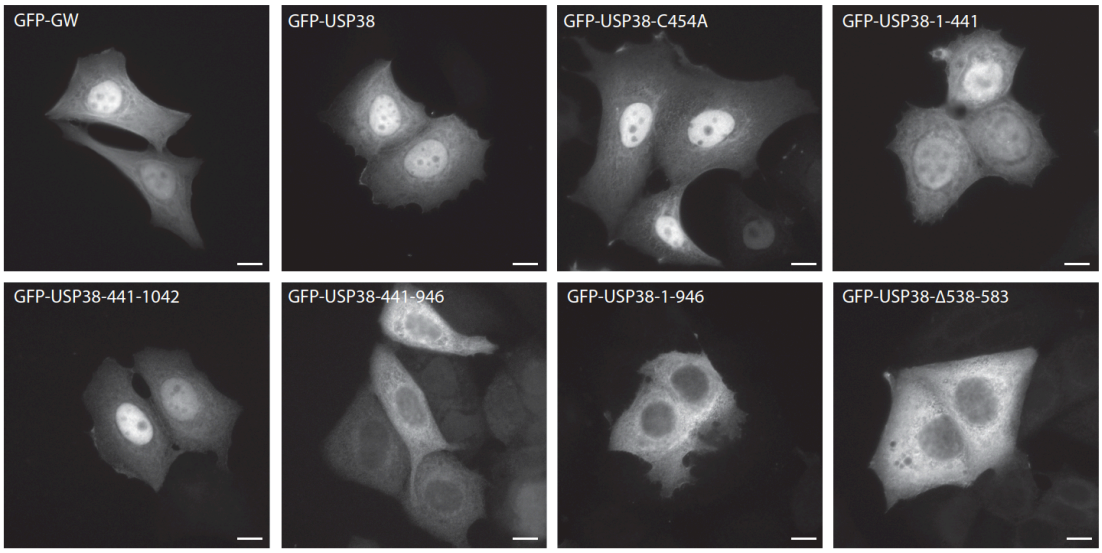

Figure 6.7 Expression level of full length and truncated USP38. MCF7 cells grown on coverslips were transfected with the indicated plasmids and were fixed using 4% PFA/PBS on the next day. Coverslips were then mounted onto mowiol plus DAPI. Images were acquired using a Nikon microscope. (Scale bar = 10µm)

Notably, these nuclear-excluded USP38 mutants all had lower expressionk levels relative to the truncation mutants USP38-1-441 and USP38-441-1042, which were nuclear localized.

6.3 Discussion

USP38 has a relatively simple domain architecture. Other than the USP domain, there is no other identifiable domain structure. Sequence alignment analysis reveals that the regions outside the USP38 catalytic domain (amino acid residues 1-440 and 947-1042) are highly conserved across the species, indicating evolutionary and functional significance of these regions. The more variable regions of USP38 correspond to the insertions between the 6 boxes of conserved residues in the catalytic USP domain as defined by Ye and colleagues (Ye *et al.*, 2009). This suggests that these insertions occured at a later stage of evolution. Such insertion of amino acid sequences between the conserved boxes of USP domain is very common among the USPs and some insertions contain independently folded domains with specific functions (Ye *et al.*, 2009). For instance, a crystallography study of the USP domain of CYLD reveals that the insertion between Box 4 and Box 5 folds into a B-box domain, which coordinates 2 zinc ions and this insertion is important for nuclear exclusion of CYLD (Komander *et al.*, 2008). By applying consensus fold recognition, Zhu reported a large number of ubiquitin like (Ubl) domains among the USPs, some of which are found in the catalytic USP domain (Zhu *et al.*, 2007). The USPs with Ubl domains in their catalytic domain are USP4, USP6, USP11, USP15, USP19, USP31, USP32 and USP43. These Ubl domains comprise of 87-89 amino acids and are found in the insertion between Box 3 and Box 4 (Ye *et al.*, 2009). This is however not the case for insertion in USP38 and there is no predicted structure for the insertions in USP38.

Localisation of a protein can sometimes provide clue to its function. For example, localisation analysis of USP21 reveals its regulatory role in centrosome and microtubule-related functions (Urbé *et al.*, 2012). However, this is not the case for USP38 as it does not show specific localisation to a particular cellular compartment or site, and is found localized to the nucleus and cytoplasm. By creating truncation mutants of USP38, I discovered that the insertion between box 2 and box 3 of the catalytic domain of USP38 is required for its nuclear localisation, since the deletion of this region leads to its nuclear exclusion (Figure 6.6). Since a consensus nuclear

localisation motif K-K/R-X/K/R is not found in this insertion, it is possible that this region associates with other proteins, which is responsible for the nuclear localisation of USP38 or perhaps the region contains a novel nuclear localisation motif. This is not the first instance of an insertion within the catalytic domain affecting localisation of a DUB, a USP in particular. For instance, the B-box found within the USP domain of CYLD is important for its nuclear exclusion or cytoplasmic retention, since the deletion of the B-Box domain leads to nuclear localisation of the mutant protein (Komander *et al.*, 2008). The localisation of USP33 is also influenced by the insertion in its catalytic domain (Thorne *et al.*, 2011). The insertion between box 3 and box 4 of the catalytic domain of USP33 contains a sequence shared by all three USP33 isoforms, which is important for ER-association. Interestingly, deletion of a specific 8 amino acids within this insertion in USP33 isoform 3 leads to its exclusive Golgi localisation. It will be interesting to assess the effect of deletion of other insertions within the USP38 catalytic domain on localisation of the protein.

Another region required for nuclear localisation of USP38 is the ~100 amino acids C-terminal to the catalytic domain (947-1042) since the deletion of this region renders the protein nuclear excluded (Figure 6.6). The localisation of USP38-1-441 in the nucleus however contradicts this statement. One possibility is that the USP38-1-441 protein is relatively small that it can shuttle freely between cytoplasm and nucleus. Moreover, the USP38-1-441 truncation mutant is not totally excluded from the nucleoli, unlike the full length USP38 and the 441-1042 truncation mutants, potentially indicating a non-specific localisation of the mutant protein.

The last 100 amino acids of USP38 also does not contain a consensus nuclear localisation motif and therefore possibly mediates interaction with other proteins, which assist in translocation of USP38 into the nucleus. The fact that the deletion of either the insertion or the last ~100 amino acids prevents nuclear localisation of USP38 suggests that these 2 stretches of amino acids are not functionally redundant and are possibly important for regulatory role of USP38 in different processes. Besides, the nuclear localisation of USP38 is perhaps important for the regulation of the protein itself, since the nuclear localised truncation mutants (USP38-441-1042

and USP38-1-441) are expressed at higher level and the nuclear excluded truncation mutants (USP38-1-946, USP38-441-946, USP38-Δ538-583) are expressed at lower level (Figure 6.5 & 6.6).

I have also revealed differential poly-ubiquitylation status of the different USP38 proteins. The catalytically inactive mutant of USP38 is more polyubiquitylated than the wildtype USP38. This is typical to catalytic mutant of DUB, as has also been demonstrated for BAP1 in Chapter 4 (Figure 4.16). However, intriguingly, despite the higher level of polyubiquitylation, the C454A mutant actually ran at lower molecular weight than the wildtype (Figure 6.6 GFP blot). The band observed on the GFP blot therefore represents a non-ubiquitin modified species of USP38 and the slight difference in molecular weight could be due to difference in phosphorylation status of the wildtype and mutant USP38. All the nuclear excluded USP38 truncation mutants were modified by polyubiquitylation, among which USP38-Δ538-583 was most heavly polyubiquitylated. This is probably because the deletion of the insertion sequence also abolished the catalytic activity of the USP domain. The USP38-441-1042 mutant showed no detectable polyubiquitylation, but this is most likely not due to loss of ubiquitylation site since the USP38-441-946 mutant was polyubiquitylated as well. Rather, such difference in ubiquitylation status of the two mutants is likely attributable to the regulation of the protein which is dependent on its nuclear localisation.

6.4 Conclusion

By creating truncation mutants of USP38, I revealed the insertion between box 2 and box 3 of USP domain of USP38 and the last 100 amino acids on C-terminus are required for nuclear localisation of USP38. As the functional significance of the different insertions is unknown, it will be interesting to check the effect of deletion of these insertions on USP38 localisation. From the data presented, it is likely that the regulation of USP38 itself is dependent on its nuclear localisation as the truncation mutants exhibiting nuclear localisation has higher expression level while those which are nuclear excluded has lower expression level. To annotate function of USP38, one potential approach is to identify its interacting partners by mass spectrometry or Y2H. Alternatively, global transcriptional analysis by RNA sequencing or DNA microarray following USP38 siRNA depletion can help identify cellular processes which regulated by this DUB.

Chapter 7

Conclusion

This chapter aims to summarise the findings reported in this thesis and to outline future perspectives. There are two main strands of this PhD research project, which are to identify DUBs involed in regulation of the AJ components, namely E-cadherin and β-catenin.

It was found in MCF7 cells that E-cadherin is constitutively degraded via the lysosomal pathway. This observation is unique to MCF7 cells, amongst cell lines (DU145 and A549 cells) tested, suggesting different regulatory mechanisms of E-cadherin at work in different cell types. Although it is not within the scope of the research aims of this project, it would be interesting to understand the underlying genetic programme that dictates such mode of E-cadherin regulation and the biological significance of this mechanism to MCF7 cells. Stemming from the constitutive lysosomal turnover of E-cadherin in MCF7 cells, a siRNA DUB library screen was performed and USP38 was identified as a leading target involved in regulation of E-cadherin lysosomal turnover. Its depletion resulted in a significant increase in the full length to 80kDa E-cadherin ratio. Notably, the loss of USP38 also resulted in a decrease in total E-cadherin level but the cells still retained adhesive properties. This loss of E-cadherin at the protein level was not accompanied by a decrease in E-cadherin mRNA level and was not due to enhanced protein turnover, indicating USP38 does not regulate E-cadherin at the transcriptional or the post-translational level. The observed loss of E-cadherin was therefore likely due to a defect in its translation. Otherwise, there could be a more intricate regulatory network involving USP38 and E-cadherin that the experimental results obtained herein cannot be interpreted straightforwardly. One indication of this came from an

observation, obtained at late stage of the research project, that the siRNA depletion of USP38 resulted in a more dramatic decrease in β-catenin level.

For future work, a rescue experiment needs to be done to definitively confirm the functional relationship between E-cadherin and USP38. Since USP38 knockdown also leads to loss of β-catenin, it is necessary to check if they are functionally related and also to assess the effect of USP38 knockdown and/or overexpression on other AJ components or known regulators of the AJ. Besides, experiments need to be done to determine the integrity and/or functionality of the translational machinery after USP38 depletion, since the ribosomal proteins, RPL7 and RPS12 are found co-immunoprecipitated with USP38 (Sowa *et al.* , 2009).

In parallel to the E-cadherin related work, BAP1 was identified as a positive regulator of β-catenin, since silencing of BAP1 expression resulted in loss of β-catenin. Subsequent work demonstrated that BAP1 regulates β-catenin at the transcriptional level. BAP1 is now a well-established tumour suppressor and has been found mutated in a variety of cancers (see section 1.2.3.6.1). On the other hand, β-catenin is a proto-oncogene, whose aberrant stabilisation is implicated in many cancers (see section 1.1.2.2.3). Thus, this observed functional relationship between a key tumour suppressor and a potent oncogene is of great research interest and may hold potential therapeutic values. Future works should definitely investigate the mechanism by which BAP1 regulates β-catenin expression. The first step towards this goal is to assess the knockdown effect of HCF1 and other transcriptional proteins (co-immunoprecipitated with BAP1) on β-catenin expression. BAP1 is known to associate with transcriptionally active chromatin (Machida *et al.* , 2009; Yu *et al.* , 2010). Therefore, a chromatin-immunoprecipitation (ChIP) assay should be performed to determine (by PCR) the presence of the promoter or regulatory sequence of the *CTNNB1* gene encoding β-catenin. BAP1 specific inhibitors offer an alternative strategy to assess BAP1-dependence of β-catenin expression. From a therapeutic point of view, it is interesting to know the effect of BAP1 inhibition in cancers arising from aberrant β-catenin stabilisation. It is

also of equal importance to assess effect of BAP1 inhibition in normal cells, since the inhibition of a tumour suppressor may have tumorigenic consequences. Besides, the regulatory role of BAP1 in the context of Wnt signaling was not investigated and this is another outstanding question to be addressed.

Stemming from the interest to identify DUBs regulating the AJ components, additional research effort was directed towards characterising two DUBs, namely USP38 and BAP1.

USP38 is a DUB with unknown function and has no other predicted domain except for its USP domain. Preliminary characterisation work showed that USP38 localises to cytoplasm and nucleus. The nuclear localisation of USP38 is dictated by an insertion of 46 aa in its USP domain and the C-terminal 100 aa. Of note, these two regions are functionally redundant and therefore may regulate distinct nuclear localisation-dependent activities of USP38. The particular finding in relation to the insertion in the USP38 USP domain provides yet another instance of amino acid insertion between the conserved boxes of the catalytic domain representing a determinant of protein localisation. The functional roles of USP38 have not been thoroughly investigated. In this regard, for future functional annotation of USP38, one approach is to perform gene expression profiling following USP38 siRNA depletion. Bioinformatic analysis of USP38, to determine potential functional motif, such as phosphorylation motif, combined with cellular or biochemical studies, will certainly facilitate elucidation of USP38 function.

On the other hand, the rescue expriment and preliminary works to characterise BAP1 have also led to some interesting observations, which are mainly associated with the BAP1-A95D mutant: The expression level of this mutant is dependent on the presence of endogenous BAP1 and it is not polyubiquitylated, unlike the wildtype and C91S mutant. Moreover, this mutant exhibits a different subcellular distribution. The study of this mutant may provide important insights into the biology of BAP1. One of the first questions to be answered is whether the A95D mutant is able to form complex with the known interacting partners, such as HCF1

and YY1. It is also interesting to know whether the ubiquitylation status of BAP1 affects its activity and association with other proteins or not. A SILAC (stable isotope labeling by amino acid in cell culture) mass spectrometry experiment to compare immunocomplexes associated with wildtype BAP1 and the A95D mutant may reveal unique interacting partners. In addition, mass spectrometry experiment can be done to identify cytoplasmic binding partners of BAP1, since the proteins found in complex with BAP1 so far are mostly nuclear transcriptional proteins (Machida *et al.* , 2009; Yu *et al.* , 2010).

To complement the human DUB siRNA library that is available in our laboratory, a human DUB esiRNA library was developed. Quality control experiments upon a sample of selected DUBs have confirmed the ability of esiRNAs to induce silencing of target genes, although with varying efficacy. The use of an alternative region within the transcript for esiRNA production may improve knockdown efficiency in these cases. Additionally, experiment conditions for efficient knockdown in various cell lines were optimised. The generation of the esiRNAs allows an alternative screening strategy, which arguably leads to less false positive hits (i.e. less off-target effects) and can therefore speed up the process of target identification. Moreover, it offers an extra layer of validation to siRNA screen experiments. Indeed, use of esiRNA against BAP1 has validated the downregulatory effect of siRNA depletion of BAP1 on β-catenin in MCF7 cells. In addition, the use of esiRNA is highly cost-effective and this is especially desirable for large-scale experiments such as mass spectrometry experiments.

References

Aberle, H., Bauer, A., Stappert, J., Kispert, A. and Kemler, R. (1997). β-catenin is a target for the ubiquitin-proteasome pathway. *EMBO J* **16**(13): 3797-804.

Aberle, H., Butz, S., Stappert, J., Weissig, H., Kemler, R. and Hoschuetzky, H. (1994). Assembly of the cadherin-catenin complex in vitro with recombinant proteins. *J Cell Sci* **107(Pt 12)**: 3655-63.

Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422**: 198-207.

Agromayor, M. and Martin-Serrano, J. (2006). Interaction of AMSH with ESCRT-III and deubiquitination of endosomal cargo. *J Biol Chem* **281**(32): 23083-91.

Ai, L., Tao, Q., Zhong, S., Fields, C. R., Kim, W. J., Lee, M. W., Cui, Y., Brown, K. D. and Robertson, K. D. (2006). Inactivation of Wnt inhibitory factor-1 (WIF1) expression by epigenetic silencing is a common event in breast cancer. *Carcinogenesis* **27**(7): 1341-8.

Aksamitiene, E., Hoek, J. B., Kholodenko, B. and Kiyatkin, A. (2007) Multistrip Western blotting to increase quantitative data output. *Electrophoresis* **28**: 3163 – 3173.

Al-Hakim, A. K., Zagorska, A., Chapman, L., Deak, M., Peggie, M. and Alessi, D. R. (2008). Control of AMPK-related kinases by USP9X and atypical Lys(29)/Lys(33)-linked polyubiquitin chains. *Biochem J* **411**(2): 249-60.

Amarzguioui, M., Holen, T., Babaie, E. and Prydz, H. (2003). Tolerance for mutations and chemical modifications in a siRNA. *Nucleic Acids Res* **31**, 589-595.

Angers, S., Thorpe, C. J., Biechele, T. L., Goldenberg, S. J., Zheng, N., MacCoss, M. J. and Moon, R. T. (2006). The KLHL12-Cullin-3 ubiquitin ligase negatively regulates the Wnt-β-catenin pathway by targeting Dishevelled for degradation. *Nat Cell Biol* **8**(4): 348-57.

Angst, B. D., Marcozzi, C., and Magee, A. I. (2001). The cadherin superfamily: diversity in form and function. *J Cell Sci* **114**(Pt 4): 629-41.

Archer, C. T., Delahodde, A., Gonzalez, F., Johnston, S. A. and Kodadek, T. (2008). Activation domain-dependent monoubiquitylation of Gal4 protein is essential for promoter binding in vivo. *J Biol Chem* **283**(18): 12614-23.

Arrigo, A., Simon, S., Gibert, B., Kretz-Remy, C., Nivon, M., Czekalla, A., Guillet, D., Moulin, M., Diaz-Latoud, C. and Vicart, P. (2007) Hsp27 (HspB1) and αB-cystallin (HspB5) as therapeutic targets. *FEBS Letters* **581**: 3665-3674.

Aza-Blanc P., Cooper, C. L., Wagner, K., Batalov, S., Deveraux, Q.L. and Cooke, M.P. (2003). Identification of modulators of TRAIL-induced apoptosis via RNAi-based phenotypic screening. *Mol. Cell* **12**: 627–37.

Baboshina, O. V. and Haas, A. L. (1996). Novel multiubiquitin chain linkages catalyzed by the conjugating enzymes E2EPF and RAD6 are recognized by 26 S proteasome subunit 5. *J Biol Chem* **271**(5): 2823-31.

Baranwal, S. and S. K. Alahari (2009). Molecular mechanisms controlling E-cadherin expression in breast cancer. *Biochem Biophys Res Commun* **384**(1): 6-11.

Baker, R. T. and Board, P. G. (1991). The human ubiquitin-52 amino acid fusion protein gene shares several structural features with mammalian ribosomal protein genes. *Nucleic Acids Res* **19**(5): 1035-40.

Balakirev, M. Y., Tcherniuk, S. O., Jaquinod, M. and Chroboczek, J. (2003). Otubains: a new family of cysteine proteases in the ubiquitin pathway. *EMBO Rep* **4**(5): 517-22.

Barker, N. and H. Clevers (2006). Mining the Wnt pathway for cancer therapeutics. *Nat Rev Drug Discov* **5**(12): 997-1014.

Batlle, E., Sancho, E., Francí, C., Domíguez, D., Monfar, M., Baulida, J. and García De Herreros, A. (2000). The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat Cell Biol* **2**(2): 84-9.

Batra, S., Shi, Y., Kuchenberker, K. M., He, B., Reguart, N., Mikami, N., You, L., Xu, Z., Lin, Y. C., Clément, G., Jablons, D. M. (2006). Wnt inhibitory factor-1, a Wnt antagonist, is silenced by promoter hypermethylation in malignant pleural mesothelioma. *Biochem Biophys Res Commun* **342**(4): 1228-32.

Batsche, E., Muchardt, C., Behrens, J., Hurst, H. C. and Crémisi, C. (1998). RB and c-Myc activate expression of the E-cadherin gene in epithelial cells through interaction with transcription factor AP-2. *Mol Cell Biol* **18**(7): 3647-58.

Behrens, J., Löwrick, O., Klein-Hitpass, L. and Birchmeier, W. (1991). The E-cadherin promoter: functional analysis of a G.C-rich region and an epithelial cell-specific palindromic regulatory element. *Proc Natl Acad Sci U S A* **88**(24): 11495-9.

Behrens, J., Mareel, M. M., Van Roy, F. M. and Birchmeier, W. (1989). Dissecting tumor cell invasion: epithelial cells acquire invasive properties after the loss of uvomorulin-mediated cell-cell adhesion. *J Cell Biol* **108**(6): 2435-47.

Behrens, J., von Kries, J. P., Kühl, M., Bruhn, L., Wedlich, D., Grosschedl, R. and Birchmeier, W. (1996). Functional interaction of ß-catenin with the transcription factor LEF-1. *Nature* **382**(6592): 638-42.

Ben-Saadon, R., Zaaroor, D., Ziv, T. and Ciechanover, A. (2006). The polycomb protein Ring1B generates self atypical mixed ubiquitin chains required for its in vitro histone H2A ligase activity. *Mol Cell* **24**(5): 701-11.

Berns K, Hijmans E.M., Mullenders, J., Brummelkamp, T. R., Velds, A., Heimerikx, M., Kerkhoven, R. M., Madiredjo, M., Nijkamp, W., Weigelt, B., Agami, R., Ge, W., Cavet, G., Linsley, P. S., Beijersbergen, R. L. and Bernards, R. (2004). A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* **428**: 431–7.

Bilić, J., Huang, Y. L., Davidson, G., Zimmermann, T., Cruciat, C. M., Bienz, M. and Niehrs, C. (2007). Wnt induces LRP6 signalosomes and promotes dishevelled-dependent LRP6 phosphorylation. *Science* **316**(5831): 1619-22.

Birchmeier, W. (2005). Cell adhesion and signal transduction in cancer. Conference on cadherins, catenins and cancer. *EMBO Rep* **6**(5): 413-7.

Birchmeier, W. and J. Behrens (1994). Cadherin expression in carcinomas: role in the formation of cell junctions and the prevention of invasiveness. *Biochim Biophys Acta* **1198**(1): 11-26.

Blaschuk, O. W., Sullivan, R., David, S. and Pouliot, Y. (1990). Identification of a cadherin cell adhesion recognition sequence. *Dev Biol* **139**(1): 227-9.

Bolós, V., Peinado, H., Pérez- Moreno, M. A., Fraga, M. F., Esteller, M. and Cano, A. (2003). The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. *J Cell Sci* **116**(Pt 3): 499-511.

Bonazzi, M., Veiga, E., Pizarro-Cerda, J. and Cossart, P. (2008) Successive post-translational modifications of E-cadherin are required for InlA-mediated internalization of Listeria monocytogenes. *Cellular Microbiology* **10**: 2208-2222.

Borden, K. L. (2000). RING domains: master builders of molecular scaffolds? *J Mol Biol* **295**(5): 1103-12.

Borodovsky, A., Kessler, B. M. , Casagrande, R., Overkleeft, H.S., Wilkinson, K.D. and Ploegh, H.L. (2001). "A novel active site-directed probe specific for

deubiquitylating enzymes reveals proteasome association of USP14." *EMBO J* **20**(18): 5187-96.

Bott, M., Brevet, M., Taylor, B. S., Shimizu, S., Ito, T., Wang, L., Creaney, J., Lake, R. A., Zakowski, M. F., Reva, B., Sander, C., Delsite, R., Powell, S., Zhou, Q., Shen, R., Olshen, A., Rusch, V. and Ladanyi, M. (2011). The nuclear deubiquitinase BAP1 is commonly inactivated by somatic mutations and 3p21.1 losses in malignant pleural mesothelioma. *Nat Genet* **43**(7): 668-72.

Bracke, M. E., Van Roy, F. M. and Mareel, M. M. (1996). The E-cadherin/catenin complex in invasion and metastasis. *Curr Top Microbiol Immunol* **213 (Pt 1)**: 123-61.

Brembeck, F. H., Rosário, M. and Birchmeier, W. (2006). Balancing cell adhesion and Wnt signaling, the key role of ß-catenin. *Curr Opin Genet Dev* **16**(1): 51-9.

Brembeck, F. H., Schwarz-Romond, T., Bakkers, J., Wilhelm, S. Hammerschmidt, M. and Birchmeier, W. (2004). Essential role of BCL9-2 in the switch between ß-catenin's adhesive and transcriptional functions. *Genes Dev* **18**(18): 2225-30.

Bremm, A., Freund, S. M. V. and Komander, D. (2010). Lys11-linked ubiquitin chains adopt compact conformations and are preferentially dydrolysed by the deubiquitinase Cezanne. *Natural Structural & Molecular Biology* **17**: 939-947.

Bryant, D. M., Kerr, M. C., Hammond, L. A., Joseph, S. R., Mostov, K. E., Teasdale, R. D. and Stow, J. L. (2007). EGF induces macropinocytosis and SNX1-modulated recycling of E-cadherin. *J Cell Sci* **120**(Pt 10): 1818-28.

Bryant, D. M. and J. L. Stow (2004). The ins and outs of E-cadherin trafficking. *Trends Cell Biol* **14**(8): 427-34.

Burnett, W. N. (1981) "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and

radiographic detection with antibody and radioiodinated protein A. *Anal Biochem* **112**, 195–203.

Caldwell, G. M., Jones, C., Gensberg, K., Jan, S., Hardy, R. G., Byrd, P., Chughtai, S., Wallis, Y., Matthews, G. M. and Morton, D. G. (2004). The Wnt antagonist sFRP1 in colorectal tumorigenesis. *Cancer Res* **64**(3): 883-8.

Cano, A., Pérez-Moreno, M. A., Rodrigo, I., Locascio, A, Blanco, M. J., del Barrio, M. G., Portillo, F. and Nieto, M. A. (2000). The transcription factor Snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat Cell Biol* **2**(2): 76-83.

Caplen, N. J., Parrish, S., Imani, F., Fire, A. and Morgan, R. A. (2001). Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc Natl Acad Sci USA* **98**, 9742-9747.

Carlsson, S. R., Roth, J., Piller, F. and Fukuda, M. (1988). Isolation and characterization of human lysosomal membrane glycoproteins, h-lamp1 and h-lamp2. Major sialoglycoproteins carrying polylactosaminoglycan. *Journal of Biochemistry* **263**: 18911 - 18919.

Catic, A., Fiebiger, E., Korbel, G. A., Blom, D., Galardy, P. J. and Ploegh, H. L. (2007). Screen for ISG15-crossreactive Deubiquitinases. *PLos One* **2**(7): 679.

Chastagner, P., Israël, A., Brou, C. (2006). Itch/AIP4 mediates Deltex degradation through the formation of K29-linked polyubiquitin chains. *EMBO Rep* **7**(11): 1147-53.

Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D. J., Gonda, D. K. and Varshavsky, A. (1989). A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* **243**(4898): 1576-83.

Chen, Y. T., Stewart, D. B. and Nelson, W. J. (1999). Coupling assembly of the E-cadherin/beta-catenin complex to efficient endoplasmic reticulum exit and basal-

lateral membrane targeting of E-cadherin in polarized MDCK cells. *J Cell Biol* **144**(4): 687-99.

Cheng, C. W., Wu, P. E., Yu, J. C., Huang, C. S., Yue, C. T., Wu, C. W. and Shen, C. Y.(2001). Mechanisms of inactivation of E-cadherin in breast carcinoma: modification of the two-hit hypothesis of tumor suppressor gene. *Oncogene* **20**(29): 3814-23.

Chi, J. T., Chang, H. Y., Wang, N. N., Chang, D. S., Dunphy, N. and Brown, P. O. (2003). Genomewide view of gene silencing by small interfering RNAs. *Proc Natl Acad Sci USA* **100**(11): 6343-6346.

Chitalia, V. C., Foy, R. L., Bachschmid, M. M., Zeng, L., Panchenko, M. V., Zhou, M. I., Bharti, A., Seldin, D. C., Lecker, S. H., Dominguez, I. and Cohen, H. T. (2008). Jade-1 inhibits Wnt signalling by ubiquitylating β-catenin and mediates Wnt pathway inhibition by pVHL. Nat Cell Biol **10**(10): 1208-16.

Choi, H. J., Huber, A. H. and Weis, W. I.(2006). Thermodynamics of ß-catenin-ligand interactions: the roles of the N- and C-terminal tails in modulating binding affinity. *J Biol Chem* **281**(2): 1027-38.

Choi, J., Park, S. Y., Costantini, F., Jho, E. H. and Joo, C. K. (2004). Adenomatous polyposis coli is down-regulated by the ubiquitin-proteasome pathway in a process facilitated by Axin. *J Biol Chem* **279**(47): 49188-98.

Ciechanover, A. (2005). Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nat Rev Mol Cell Biol* **6**(1): 79-87.

Ciechanover, A., Elias, S., Heller, H. and Hershko, A. (1982). "Covalent affinity" purification of ubiquitin-activating enzyme. *J Biol Chem* **257**(5): 2537-42.

Ciechanover, A., Heller, H., Elias, S., Haas, A. L. and Hershko, A. (1980). ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. *Proc Natl Acad Sci U S A* **77**(3): 1365-8.

Ciechanover, A., Hod, Y. and Hershko, A. (1978) A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochem. Biophys. Res Commun.* **81**: 1100-1105

Clague, M. J. and Urbé, S. (2010) Ubiquitin: Same Molecule, Different Degradation Pathways. *Cell* **143** (5): 682-685.

Clague, M. J., Liu, H. and Urbé, S. (2012) Governance of endocytic trafficking and signaling by reversible ubiquitylation. *Dev. Cell.* **23**: 457 – 467.

Clevers, H. (2004). Wnt breakers in colon cancer. *Cancer Cell* **5**(1): 5-6.

Clevers, H. (2006). Wnt/ß-catenin signaling in development and disease. *Cell* **127**(3): 469-80.

Comijn, J., Berx, G., Vermassen, P., Verschueren, K., van Grunsven, L., Bruyneel, E., Mareel, M., Huylebroeck, D. and van Roy, F. (2001). The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Mol Cell* **7**(6): 1267-78.

Cooper, E. M., Cutcliffe, C., Kristiansen, T. Z., Pandey, A., Pickart, C. M. and Cohen, R. E. (2009). K63-specific deubiquitination by two JAMM/MPN+ complexes: BRISC-associated Brcc36 and proteasomal Poh1. *EMBO J* **28**(6): 621-31.

Cope, G. A., Suh, G. S., Aravind, L., Schwarz, S. E., Zipursky, S. L., Koonin, E. V. and Deshaies, R. J. (2002). Role of predicted metalloprotease motif of Jab1/Csn5 in cleavage of Nedd8 from Cul1. *Science* **298**(5593): 608-11.

Cullen, B. R. (2006). Enhancing and confirming the specificity of RNAi experiments. *Nature Methods* **3**, 677-681.

Cullen, L. M. and Arndt, G. M. (2005). Genome-wide screening for gene function using RNAi in mammalian cells. *Immunology and Cell Biology* **83**, 217-223.

Cummins, J. M., Rago, C., Kohli, M., Kinzler, K. W., Lengauer, C. and Vogelstein, B. (2004a). Tumour suppression: disruption of HAUSP gene stabilizes p53. *Nature* **428**(6982): 1 p following 486.

Cummins, J. M. and Vogelstein, B. (2004). HAUSP is required for p53 destabilization. *Cell Cycle* **3**(6): 689-92.

D'Souza-Schorey, C. (2005). Disassembling adherens junctions: breaking up is hard to do. *Trends Cell Biol* **15**(1): 19-26.

Daugherty, R. L. and C. J. Gottardi (2007). Phospho-regulation of Beta-catenin adhesion and signaling functions. *Physiology (Bethesda)* **22**: 303-9.

Davidson, G., Wu, W., Shen, J., Bilic, J., Fenger, U., Stannek, P., Glinka, A. and Niehrs, C. (2005). Casein kinase 1 ɣ couples Wnt receptor activation to cytoplasmic signal transduction. *Nature* **438**(7069): 867-72.

Davis, M. A., Ireton, R. C. and Reynolds, A. B. (2003). A core function for p120-catenin in cadherin turnover. *J Cell Biol* **163**(3): 525-34.

de Beco, S., Gueudry, C., Amblard, F. and Coscoy, S. (2009). Endocytosis is required for E-cadherin redistribution at mature adherens junctions. *Proc Natl Acad Sci U S A* **106**(17): 7010-5.

De Duve, C., Gianetto, R., Applemans, F. and Wattiaux, R. (1953). Enzymic content of the mitochondria fraction. *Nature* **172**(4390): 1143-4.

De Wever, O., Derycke, L., Hendrix, A., De Meerleer, G., Godeau, F., Depypere, H. and Bracke, M. (2007). Soluble cadherins as cancer biomarkers. *Clin Exp Metastasis* **24**(8): 685-97.

de Rooji, J., Kerstens, A., Danuser, G.Schwartz, M. A. and Waterman-Storer, C. M. (2005). Intergrin-dependent actomyosin contraction regulates epithelial cell scattering. *J Cell Biol* **171**(1) 153-164.

Decary, S., Decesse, J. T., Ogryzko, V., Reed, J. C., Naguibneva, I. Harel-Bellan, A. and Cremisi, C. E. (2002). The retinoblastoma protein binds the promoter of the survival gene bcl-2 and regulates its transcription in epithelial cells through transcription factor AP-2. *Mol Cell Biol* **22**(22): 7877-88.

Deshaies, R. J. and Joazeiro, C. A. (2009). RING domain E3 ubiquitin ligases. *Annu Rev Biochem* **78**: 399-434.

Dey, A., Seshasayee, D., Noubade, R., French, D. M., Liu, J., Chaurushiya, M. S., Kirkpatrick, D. S., Pham, V. C., Lill, J. R., Bakalarski, C. E., Wu, J., Phu, L., Katavolos, P., LaFave, L. M., Abdel-Wahab, O., Modrusan, Z., Seshagirl, S., Dong, K., Lin, Z., Balazs, M., Suriben, R., Newton, K., Hymowitz, S., Garcia-Manero, G., Martin, F., Levine, R. L. and Dixit, V. M.(2012). Loss of the tumor suppressor BAP1 causes myeloid transformation. *Science* **337**(6101): 1541-6.

Dikic, I., Wakatsuki, S. and Walters, K. J. (2009). Ubiquitin-binding domains - from structures to functions. *Nat Rev Mol Cell Biol* **10**(10): 659-71.

Dixit, V. M., Green, S., Sarma, V., Holzman, L. B., Wolf, F. W., O'Rourke, K., Ward, P. A., Prochownik, E. V. and Marks, R. M. (1990). Tumor necrosis factor-α induction of novel gene products in human endothelial cells including a macrophage-specific chemotaxin. *J Biol Chem* **265**(5): 2973-8.

Dong, Y., Hakimi, M. A., Chen, X., Kumaraswamy, E., Cooch, N. S., Godwin, A. K. and Shiekhattar, R. (2003). Regulation of BRCC, a holoenzyme complex containing BRCA1 and BRCA2, by a signalosome-like subunit and its role in DNA repair. *Mol Cell* **12**(5): 1087-99.

Drees, F., Pokutta, S., Yamada, S., Nelson, W. J. and Weis, W. I. (2005). Alpha-catenin is a molecular switch that binds E-cadherin-beta-catenin and regulates actin-filament assembly. *Cell* **123**(5): 903-15.

Driscoll, J. and Goldberg, A. L. (1990). The proteasome (multicatalytic protease) is a component of the 1500-kDa proteolytic complex which degrades ubiquitin-conjugated proteins. *J Biol Chem* **265**(9): 4789-92.

Duncan, L. M., Piper, S., Dodd, R. B., Saville, M. K., Sanderson, C. M., Luzio, J. P. and Lehner, P. J. (2006). Lysine-63-linked ubiquitination is required for endolysosomal degradation of class I molecules. *EMBO J* **25**(8): 1635-45.

Dupre-Crochet, S., Figueroa, A., Hogan, C., Ferber, E. C., Bialucha, C. U., Adams, J., Richardson, E. C. N. and Fujita, Y. (2007). Casein kinase 1 is a novel negative regulator of E-cadherin-based cell-cell contacts. *Mol Cell Biol* **27**(10): 3804-16.

Dye, B. T. and Schulman, B. A. (2007). Structural Mechanisms Underlying Posttranslational Modification by Ubiquitiin-Like Proteins. *Annual Review of Biophysics and Biomolecular Structure* **36**, 131-150.

Echeverri, C. J., Beachy, P. A., Baum, B., Boutros, M., Buchholz, F., Chanda, S. K., Downward, J., Ellenberg, J., Fraser, A. G., Hacohen, N., Hahn, W. C., Jackson, A. L., Kiger, A., Linsley, P. S., Lum, L., Ma, Y., Mathey-Prévôt, B., Root, D. E., Sabatini, D. M., Taipale, J., Perrimon, N. and Bernards, R. (2006). Minimising the risk of reportung false positives in large-scale RNAi screens. *Nature Methods* **3**, 777-779.

Edelmann, M. J., Iphöfer, A., Akutsu, M., Altun, M, di Gleria, K., Kramer, H. B., Fiebiger, E., Dhe-Paganon, S. and Kessler, B. M. (2009). Structural basis and specificity of human otubain 1-mediated deubiquitination. *Biochem J* **418**(2): 379-90.

Eger, A., Aigner, K., Sonderegger, S., Dampier, B., Oehler, S., Schreiber, M., Berx, G., Cano, A., Beug, H. and Foisner, R. (2005). DeltaEF1 is a transcriptional

repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. *Oncogene* **24**(14): 2375-85.

Elabshir, S. M., Marinez, J., Patkaniowska, A., Lendeckel, W. and Tushl, T. (2002). Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. *EMBO J* **20**, 6877-6888.

Eletr, Z. M. and Wilkinson, K. D. (2011). An emerging model for BAP1's role in regulating cell cycle progression. *Cell Biochem Biophys* **60**(1-2): 3-11.

Elsasser, S., Chandler-Militello, D., Müller, B., Hanna, J. and Finley, D. (2004). Rad23 and Rpn10 serve as alternative ubiquitin receptors for the proteasome. *J Biol Chem* **279**(26): 26817-22.

Ennis, H. L. and Lubin, M. (1964). Cycloheximide: Aspects of ihibition of protein synthesis in mammalian cells. *Science* **146**(3650): 1474-1476.

Etlinger, J. D. and Goldberg, A. L. (1977). A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes. *Proc Natl Acad Sci U S A* **74**(1): 54-8.

Fan, Y. H., Yu, Y, Mao, R. F., Tan, X. J., Xu, G. F., Zhang, H., Lu, X. B., Fu, S. B., and Yang, J. (2011). USP4 targets TAK1 to downregulate TNFα-induced NF-κB activation. *Cell Death Differ* **18**(10): 1547-60.

Farquhar, M. G. and Pallade, G. E. (1963) Junctional complexes in various epithelia. *Journal of Cell Biology* **17**: 395 – 412.

Ferber, A., Yaen, C., Sarmiento, E. and Martinez, J. (2002). An octapeptide in the juxtamembrane domain of VE-cadherin is important for p120ctn binding and cell proliferation. *Exp Cell Res* **274**(1): 35-44.

Ferro, A., Carvalho, A. L., Teixeira-Castro, A., Almeida, C., Tomé, R. J., Cortes, L., Rodrigues, A. J., Logarinho, E., Sequeiros, J., Macedo-Ribeiro, S. and Maciel, P.

(2007). NEDD8: a new ataxin-3 interactor. *Biochim Biophys Acta* **1773**(11): 1619-27.

Finley, D. (2009). Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu Rev Biochem* **78**: 477-513.

Fraile, J. M., Quesada, V., Rodríguez, D., Freije, J. M. P. and LÓpez-Otín, C. (2012). Deubiquitinases in cancer: new functions and therapeutic options. *Oncogene* **31**(19): 2373-88.

Franke, W. W. (2009) Discovering the molecular components of intracellular junctions – a historical view. *Cold Spring Harb Perspect Biol* **2**: 1 – 35.

Frixen, U. H., Behrens, J., Sachs, M., Eberle, G., Voss, B., Warda, A., Löchner, D. and Birchmeier, W. (1991). E-cadherin-mediated cell-cell adhesion prevents invasiveness of human carcinoma cells. *J Cell Biol* **113**(1): 173-85.

Fujita, Y., Krause, G., Scheffner, M., Zechner, D., Leddy, H. E. M., Behrens, J., Sommer, T. and Birchmeier, W. (2002). Hakai, a c-Cbl-like protein, ubiquitinates and induces endocytosis of the E-cadherin complex. *Nat Cell Biol* **4**(3): 222-31.

Fukazawa, T., Miyake, S., Band, V. and Band, H. (1996). Tyrosine phosphorylation of Cbl upon epidermal growth factor (EGF) stimulation and its association with EGF receptor and downstream signaling proteins. *J Biol Chem* **271**(24): 14554-9.

Galan, J. M., Moreau, V., Andre, B., Volland, C. and Haguenauer-Tsapis, R. (1996). Ubiquitination mediated by the Npi1p/Rsp5p ubiquitin-protein ligase is required for endocytosis of the yeast uracil permease. *J Biol Chem* **271**(18): 10946-52.

Geisler, S., Holmström, K. M., Skujat, D., Fiesel, F. C., Rothfuss, O. C., Kahle, P. J. and Springer, W. (2010). PINK1/Parkin-mediated mitophagy is dependent on VDAC1 and p62/SQSTM1. *Nat Cell Biol* **12**(2): 119-31.

Geng, F., Wenzel, S. and Tansey, W. P (2012). Ubiquitin and proteasomes in transcription. *Annu Rev Biochem* **81**: 177-201.

Gewies, A. and S. Grimm (2003). UBP41 is a proapoptotic ubiquitin-specific protease. *Cancer Res* **63**(3): 682-8.

Giles, R. H., van Es, J. H. and Clevers, H. (2003). Caught up in a Wnt storm: Wnt signaling in cancer. *Biochim Biophys Acta* **1653**(1): 1-24.

Gloushankova, N. A. (2008). Changes in regulation of cell-cell adhesion during tumor transformation. *Biochemistry (Mosc)* **73**(7): 742-50.

Goldberg, A. L. and St John, A. C. (1976). Intracellular protein degradation in mammalian and bacterial cells: Part 2. *Annu Rev Biochem* **45**: 747-803.

Goldknopf, I. L. and Busch, H. (1977). Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24. *Proc Natl Acad Sci U S A* **74**(3): 864-8.

Goldknopf, I. L. and Busch, H. (1975). Remarkable similarities of peptide fingerprints of histone 2A and nonhistone chromosomal protein A24. *Biochem Biophys Res Commun* **65**(3): 951-60.

Goldstein, G., Scheid, M., Hammerling, U., Schlesinger, D. H., Niall, H. D. and Boyse, E. A. (1975). Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc Natl Acad Sci U S A* **72**(1): 11-5.

Gong, L., Kamitani, T., Millas, S. and Yeh, E. T. H. (2000). Identification of a Novel Ispeptidase with Dual Specificity for Ubiquitin-and NEDD8-conjugated Proteins. *J Biol Chem* **275**, 14212-14216.

Greenburg, G. and E. D. Hay (1982). Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells. *J Cell Biol* **95**(1): 333-9.

Greenburg, G. and E. D. Hay (1986). Cytodifferentiation and tissue phenotype change during transformation of embryonic lens epithelium to mesenchyme-like cells in vitro. *Dev Biol* **115**(2): 363-79.

Greenburg, G. and E. D. Hay (1988). Cytoskeleton and thyroglobulin expression change during transformation of thyroid epithelium to mesenchyme-like cells. *Development* **102**(3): 605-22.

Greger, V., Passarge, E., Höpping, W., Messmer, E. and Horsthemke, B. (1989). Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. *Hum Genet* **83**(2): 155-8.

Groden, J., Thliveris, A., Samowitz, W., Carlson, M., Gelbert, L., Albertsen, H., Joslyn, G., Stevens, J., Spirio, L. and Robertson, M. (1991). Identification and characterization of the familial adenomatous polyposis coli gene. *Cell* **66**(3): 589-600.

Groll, M. and Clausen, T. (2003). Molecular shredders: how proteasomes fulfill their role. *Curr Opin Struct Biol* **13**(6): 665-73.

Guilford, P., Hopkins, J., Harraway, J., McLeod, M., McLeod, N., Harawira, P., Taite, H., Scoular, R., Miller, A. and Reeve, A. E. (1998). E-cadherin germline mutations in familial gastric cancer. *Nature* **392**(6674): 402-5.

Gumbiner, B., Stevenson, B. and Grimaldi, A. (1988). The role of the cell adhesion molecule uvomorulin in the formation and maintenance of the epithelial junctional complex. *J Cell Biol* **107**(4): 1575-87.

Haas, T. L., Emmerich, C. H., Gerlach, B., Schmukle, A. C., Cordier, S. M., Rieser, E., Feltham, R., Vince, J., Warnken, U., Wenger, T., Koschny, R, Komander, D., Silke, J. and Walczak, H. (2009). Recruitment of the linear ubiquitin chain assembly complex stabilizes the TNF-R1 signaling complex and is required for TNF-mediated gene induction. *Mol Cell* **36**(5): 831-44.

Haglund, K. and Dikic, I. (2005). Ubiquitylation and cell signaling. *EMBO J* **24**(19): 3353-9.

Haglund, K., Sigismund, S., Polo, S., Szymkiewicz, I., Di Fiore, P. P., Dikic, I. (2003). Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. *Nat Cell Biol* **5**(5): 461-6.

Hajra, K. M., Chen, D. Y. and Fearon, E. R. (2002). The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Res* **62**(6): 1613-8.

Hamazaki, J., Iemura, S., Natsume, T., Yashiroda, H, Tanaka, K. and Murata, S. (2006). A novel proteasome interacting protein recruits the deubiquitinating enzyme UCH37 to 26S proteasomes. *EMBO J* **25**(19): 4524-36.

Hao, H. X., Xie, Y., Zhang, Y., Charlat, O., Oster, E., Avello, M., Lei, H., Mickanin, C., Liu, D., Ruffner, H., Mao, X., Ma, Q., Zamponi, R., Bouwmeester, T., Finan, P. M., Kirshner, M. W., Porter, J. A., Serluca, F. C. and Cong, F. (2012). ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner. *Nature* **485**(7397): 195-200.

Harbour, J. W., Onken, M. D., Roberson, E. D. O., Duan, S., Cao, L., Worley, L. A., Council, M. L., Matatall, K. A., Helms, C. and Bowcock, A. M. (2010). Frequent mutation of BAP1 in metastasizing uveal melanomas. *Science* **330**(6009): 1410-3.

Hart, M., Concordet, J. P., Lassot, I., Albert, I., del los Santos, R., Durand, H., Perret, C., Rubinfeld, B., Margottin, F., Benarous, R. and Polakis, P. (1999). The F-box protein ß-TrCP associates with phosphorylated ß-catenin and regulates its activity in the cell. *Curr Biol* **9**(4): 207-10.

Hartley, J. L., Temple, G. F. and Brasch, M. A. (2000) DNA cloning using in vitro site-specific recombination. *Genome Research* **10**: 1788 – 1795.

Hartsock, A. and Nelson, W. J. (2008). Adherens and tight junctions: structure, function and connections to the actin cytoskeleton. *Biochim Biophys Acta* **1778**(3): 660-9.

Hassink, G. C., Zhao, B., Sompallae, R., Altun, M., Gastaldello, S., Zinin, N. V., Masucci, M. G. and Lindsten, K. (2009). The ER-resident ubiquitin-specific protease 19 participates in the UPR and rescues ERAD substrates. *EMBO Rep.* **10**(7): 755-761.

Hay-Koren, A., Caspi, M., Zilberberg, A. and Rosin-Arbesfeld, R. (2011). The EDD E3 ubiquitin ligase ubiquitinates and up-regulates β-catenin. *Mol Biol Cell* **22**(3): 399-411.

Henderson, B. R. and Fagotto, F. (2002). The ins and outs of APC and beta-catenin nuclear transport. *EMBO Rep* **3**(9): 834-9.

Hennig, G., Löwrick, O., Birchmeier, W. and Behrens, J. (1996). Mechanisms identified in the transcriptional control of epithelial gene expression. *J Biol Chem* **271**(1): 595-602.

Herault, Y., Michel, D., Chatelain, G. and Brun, G. (1991) cDNA and predicted amino acid sequences of the human ribosomal protein genes rpS12 and rpL17. *Nucleic Acid Res* **19**: 4001.

Herman, J. G., Latif, F., Lerman, M. I., Zbar, B., Liu, S., Samid, D., Duan, D. S., Gnarra, J. R. and Linehan, W. M. et al. (1994). Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. *Proc Natl Acad Sci U S A* **91**(21): 9700-4.

Hershko, A. and Ciechanover, A. (1998). The ubiquitin system. *Annu Rev Biochem* **67**: 425-79.

Hershko, A., Ciechanover, A., Heller, H., Haas, A. L. and Rose, I. A. (1980). Proposed role of ATP in protein breakdown: conjugation of protein with multiple

chains of the polypeptide of ATP-dependent proteolysis. *Proc Natl Acad Sci U S A* **77**(4): 1783-6.

Hershko, A., Heller, H., Elias, S. and Ciechanover, A. (1983). Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. *J Biol Chem* **258**(13): 8206-14.

Hershko, A., Leshinsky, E., Ganoth, D. and Heller, H. (1984). ATP-dependent degradation of ubiquitin-protein conjugates. *Proc Natl Acad Sci U S A* **81**(6): 1619-23.

Hicke, L. and Riezman, H. (1996). Ubiquitination of a yeast plasma membrane receptor signals its ligand-stimulated endocytosis. *Cell* **84**(2): 277-87.

Hicke, L., Schubert, H. L. and Hill, C. P. (2005). Ubiquitin-binding domains. *Nat Rev Mol Cell Biol* **6**(8): 610-21.

Higuchi, R., Dollinger, G., Walsh, P. S. and Griffith, R. (1992). Simultaneous amplification and detection of specific DNA sequences. *Biotechnology* **10**(4): 413 – 417.

Hirokawa, N. and J. E. Heuser (1981). Quick-freeze, deep-etch visualization of the cytoskeleton beneath surface differentiations of intestinal epithelial cells. *J Cell Biol* **91**(2 Pt 1): 399-409.

Hiscox, S. and Jiang, W. G. (1999). Hepatocyte growth factor/scatter factor disrupts epithelial tumour cell-cell adhesion: involvement of beta-catenin. *Anticancer Research* **19**: 509 - 517.

Hochstrasser, M. (2009). Review Article Origin and function of ubiquitin-like proteins. *Nature* **458**, 422-429.

Hoeller, D. and Dikic, I. (2009). Targeting the ubiquitin system in cancer therapy. *Nature* **458**(7237): 438-44.

Hofmann, K. and Bucher, P. (1996). The UBA domain: a sequence motif present in multiple enzyme classes of the ubiquitination pathway. *Trends Biochem Sci* **21**(5): 172-3.

Hofmann, K. and Falquet, L. (2001). A ubiquitin-interacting motif conserved in components of the proteasomal and lysosomal protein degradation systems. *Trends Biochem Sci* **26**(6): 347-50.

Holcombe, R. F., Marsh, J. L., Waterman, M. L., Lin, F., Milovanovic, T. and Truoung, T. (2002). Expression of Wnt ligands and Frizzled receptors in colonic mucosa and in colon carcinoma. *Mol Pathol* **55**(4): 220-6.

Holen, T., Amarzguioui, M., Wiiger, M. T., Babaie, E. and Prydz, H. (2002). Positional efffects of short interferin RNAs targeting the human coagulation triger Tissue Factor. *Nucleic Acids Res* **30**, 1757-1766.

Hosono, S., Gross, I., English, M. A., Hajra, K. M., Fearon, E. R., and Licht, J. D. (2000). E-cadherin is a WT1 target gene. *J Biol Chem* **275**(15): 10943-53.

Hough, R., Pratt, G. and Rechsteiner, M. (1986). Ubiquitin-lysozyme conjugates. Identification and characterization of an ATP-dependent protease from rabbit reticulocyte lysates. *J Biol Chem* **261**(5): 2400-8.

Hu, M., Li, P., Song, L. Jeffrey, P. D., Chernova, T. A., Wilkinson, K. D., Cohen, R. E. and Shi, Y. (2005). Structure and mechanisms of the proteasome-associated deubiquitinating enzyme USP14.*EMBO J* **24**(21): 3747-56.

Huang, T. T., Nijman, S. M. B., Mirchandani, K. D., Galardy, P. J., Cohn, M. A., Haas, W., Gypi, S. P., Ploegh, H. L., Bernards, R. and D'Andrea, A. D. (2006). Regulation of monoubiquitinated PCNA by DUB autocleavage. *Nat Cell Biol* **8**(4): 339-47.

Huang, X., Langelotz, C., Hetfeld-Pechoc, B. K., Schwenk, W. and Dubiel, W. (2009). The COP9 signalosome mediates β-catenin degradation by deneddylation and blocks adenomatous polyposis coli destruction via USP15. *J Mol Biol* **391**(4): 691-702.

Huber, A. H., Nelson, W. J., Weis, W. I. (1997). Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell* **90**(5): 871-82.

Huber, A. H., Stewart, D. B., Laurents, D. V., Nelson, W. J. and Weis, W. I. (2001). The cadherin cytoplasmic domain is unstructured in the absence of beta-catenin. A possible mechanism for regulating cadherin turnover. *J Biol Chem* **276**(15): 12301-9.

Huber, A. H. and W. I. Weis (2001). The structure of the beta-catenin/E-cadherin complex and the molecular basis of diverse ligand recognition by beta-catenin. *Cell* **105**(3): 391-402.

Huber, M. A., Kraut, N. and Beug, H. (2005). Molecular requirements for epithelial-mesenchymal transition during tumor progression. *Curr Opin Cell Biol* **17**(5): 548-58.

Hugo, H., Ackland, M. L., Blick, T., Lawrence, M. G., Clements, J. A., Williams, E. D. and Thompson, E. W. (2007). Epithelial--mesenchymal and mesenchymal--epithelial transitions in carcinoma progression. *J Cell Physiol* **213**(2): 374-83.

Hulsken, J., Birchmeier, W. and Behrens, J, (1994). E-cadherin and APC compete for the interaction with beta-catenin and the cytoskeleton. *J Cell Biol* **127**(6 Pt 2): 2061-9.

Hunt, L. T. and Dayhoff, M. O. (1977). Amino-terminal sequence identity of ubiquitin and the nonhistone component of nuclear protein A24. *Biochem Biophys Res Commun* **74**(2): 650-5.

Hurley, J. H. and Stenmark, H. (2011). Molecular mechanisms of ubiquitin-dependent membrane traffic. *Annu Rev Biophys* **40**: 119-42.

Hurley, J. H., Lee, S. and Prag, G. (2006). Ubiquitin-binding domains. *Biochem J* **399**(3): 361-72.

Ikeda, F., Deribe, Y. L., Skånland, S. S., Stieglitz, B., Grabbe, C., Franz-Wachtel, M., van Wijik, S. J., Goswami, P., Nagy, V., Terzic, J., Tokunaga, F., Androulidaki, A., Nakagawa, T., Pasparakis, M., Iwai, K., Sundberg, J. P., Schaefer, L., Rittinger, K., Macek, B, and Dikic, I. (2011). SHARPIN forms a linear ubiquitin ligase complex regulating NF-κB activity and apoptosis. *Nature* **471**(7340): 637-41.

Ikeda, F. and Dikic, I. (2008). Atypical ubiquitin chains: new molecular signals. 'Protein Modifications: Beyond the Usual Suspects' review series. *EMBO Rep* **9**(6): 536-42.

Ireton, R. C., Davis, M. A., van Hengel, J., Mariner, D. J., Barnes, K., Thoreson, M. A., Anastasiadis, P. Z., Matrisian, L., Bundy, L. M., Sealy, L., Gilbert., B., van Roy, F. and Reynolds, A. B. (2002). A novel role for p120 catenin in E-cadherin function. *J Cell Biol* **159**(3): 465-76.

Ito, M., Kawano, K., Miyagishi, M. and Taira, K. (2005). Genome-wide application of RNAi to the discovert of potential drug targets. *FEBS Letters* **579**(26): 5988-95.

Jackson, A. L. and Linsley, P. S. (2010). Reognising and avoiding siRNA off-target effects for target identification and therapeutic application. *Nat Rec Drug Discov* **9**(1): 57-67.

Jackson, A. L., Bartz, S. R., Schelter, J. Kobayashi, S. V., Burchard, J., Mao, M., Li, B., Cavet, G. and Linsley, P. S. (2003). Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotech* **21**(6): 635-637.

Jackson, A. L., Burchard, J., Schelter, J. Chau, B. N., Cleary, M., Lim, L. and Linsley, P. S. (2006). Widespread siRNA 'off-target' sequence complementaruty. *RNA* **12**(7): 1179-1187.

Jacobson, A. D., Zhang, N. Y., Xu, P., Han, K. J., Noone, S., Peng, J. and Liu, C. W. (2009) The Lysine 48 and Lysine 63 Ubiquitin Conjugates Are Processed Differently by the 26 S Proteasome. *Biochem J* **284,** 35485-35494.

Jensen, D. E., Proctor, M., Marquis, S. T., Gardner, H. P., Ha, S. I., Chodosh, L. A., Ishov, A. M., Tommerup, N., Vissing, H., Sekido, Y., Minna, J., Borodovsky, A., Schultz, D. C., Wilkinson, K. D., Maul, G. G., Barlev, N., Berger, S. L., Prendergast, G. C. and Rauscher 3rd, F. J. (1998). BAP1: a novel ubiquitin hydrolase which binds to the BRCA1 RING finger and enhances BRCA1-mediated cell growth suppression. *Oncogene* **16**(9): 1097-112.

Jensen, D. E. and Rauscher, 3rd, F. J. (1999). BAP1, a candidate tumor suppressor protein that interacts with BRCA1. *Ann N Y Acad Sci* **886**: 191-4.

Jin, L., Williamson, A., Banerjee, S., Philipp, I. and Rape, M. (2008). Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex. *Cell* **133**(4): 653-65.

Johnston, S. C., Larsen, C. N., Cook, W. J., Wilkinson, K. D. and Hill, C. P. (1997). Crystal structure of a deubiquitinating enzyme (human UCH-L3) at 1.8 A resolution. *EMBO J* **16**(13): 3787-96.

Johnston, S. C., Riddle, S. M., Cohen, R. E. and Hill, C. P. (1999). Structural basis for the specificity of ubiquitin C-terminal hydrolases. *EMBO J* **18**(14): 3877-87.

Kalluri, R. and Neilson, E. G. (2003). Epithelial-mesenchymal transition and its implications for fibrosis. *J Clin Invest* **112**(12): 1776-84.

Kawanishi, J., Kato, J., Sasaki, K., Fujii, S., Watanabe, N. and Niitsu, Y. (1995). Loss of E-cadherin-dependent cell-cell adhesion due to mutation of the β-catenin gene in a human cancer cell line, HSC-39. *Mol Cell Biol* **15**(3): 1175-81.

Kayagaki, N., Phung, Q., Chan, S., Chaudhari, R., Quan, C., O'Rourke, K. M., Eby, M., Pietras, E., Cheng, G., Bazan, J. F., Zhang, Z., Arnott, D. and Dixit, V. M. (2007). DUBA: a deubiquitinase that regulates type I interferon production. *Science* **318**(5856): 1628-32.

Kinzler, K. W. and Vogelstein, B. (1996). Lessons from hereditary colorectal cancer. *Cell* **87**(2): 159-70.

Kirkin, V. and Dikic, I. (2007). Role of ubiquitin- and Ubl-binding proteins in cell signaling. *Curr Opin Cell Biol* **19**(2): 199-205.

Kittler, R., Pelletier, L., Ma, C., Poser, I., Fischer, S., Hyman, A. A. and Buchholz, F. (2005). RNA interference rescue by bacterial artificial chromosome transgenesis in mammalian tissue culture cells. *Proc Natl Acad Sci USA* **102**(7): 2396-2401.

Klaus, A. and Birchmeier, W. (2008). Wnt signalling and its impact on development and cancer. Nat *Rev Cancer* **8**(5): 387-98.

Klucky, B., Mueller, R., Vogt, I., Teurich, S., Hartenstein, B., Breuhahn, K., Flechtenamarcher, (2007). Kallikrein 6 induces E-cadherin shedding and promotes cell proliferation, migration, and invasion. *Cancer Res* **67**(17): 8198-206.

Kobe, B. and Kajava, A. V. (2000). When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. *Trends Biochem Sci* **25**(10): 509-15.

Komander, D. (2009). The emerging complexity of protein ubiquitination. *Biochem Soc Trans* **37**(Pt 5): 937-53.

Komander, D. (2010) Mechanism, specificity and structure of the deubiquitinses. *Subcellular Biochemistry* **54**: 69-87.

Komander, D. and Barford, D. (2008). Structure of the A20 OTU domain and mechanistic insights into deubiquitination. *Biochem J* **409**(1): 77-85.

Komander, D. and Rape, M. (2012). The ubiquitin code. *Annu Rev Biochem* **81**: 203-29.

Komander, D., Clague, M. J. and Urbé, S. (2009). Breaking the chains: structure and function of the deubiquitinases. *Nat Rev Mol Cell Biol* **10**(8): 550-63.

Komander, D., Lord, C. J., Scheel, H., Swift, S., Hofmann, K., Ashworth, A. and Barford, D. (2008) The structure of the CYLD USP domain explains its specificity for Lys63-linked polyubiquitin and reveals a B box module. *Molecular Cell* **29**(4): 451-464.

Komander, D., Reyes-Turcu, F., Licchesi, J. D. F., Odenwaelder, P., Wilkinson, K. D. and Barford, D. (2009). Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. *EMBO Rep* **10**(5): 466-73.

Komander, D., Reyes-Turcu, F., Licchesi, J. D. F., Odenwaelder, P., Wilkinson, K. D. and Barford, D. (2009b). Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. *EMBO Rep* **10**(5): 466-73.

Koo, B. K., Spit, M., Jordens, I., Low, T. Y., Stange, D. E., van de Wetering, M., van Es, J. H., Mohammed, S., Heck, A. J., Maurice, M. M. and Clevers, H. (2012). Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. *Nature* **488**(7413): 665-9.

Lai, E. C. (2002). Micro RNAs are complementart to 3' UTR sequence motifs that mediate neagative post-transccriptional regualtion. *Nat Genet* **30**(4): 363-364.

Lake, M. W., Wuebbens, M. M., Rajapopalan, K. V. and Schindelin, H. (2001). Mechanism of ubiquitin activation revealed by the structure of a bacterial MoeB-MoaD complex. *Nature* **414**(6861): 325-9.

Lauwers, E., Jacob, C. and André, B. (2009). K63-linked ubiquitin chains as a specific signal for protein sorting into the multivesicular body pathway. *J Cell Biol* **185**(3): 493-502.

Le, T. L.,Yap, A. S. and Stow, J. L. (1999). Recycling of E-cadherin: a potential mechanism for regulating cadherin dynamics." *J Cell Biol* **146**(1): 219-32.

Lee, A. Y., He, B., You, L., Dadfarmay, S., Xu, Z., Mazieres, J., Mikami, I., McCormick, F. and Jablons, D. M. (2004). Expression of the secreted frizzled-related protein gene family is downregulated in human mesothelioma. *Oncogene* **23**(39): 6672-6.

Lee, E. G., Boone, D. L., Chai, S., LIbby, S. L., Chien, M., Lodolce, J. P. and Ma, A. (2000). Failure to regulate TNF-induced NF-κB and cell death responses in A20-deficient mice. *Science* **289**(5488): 2350-4.

Leimkühler, S., Wuebbens, M. M. and Rajagopalan, K. V. (2001). Characterization of Escherichia coli MoeB and its involvement in the activation of molybdopterin synthase for the biosynthesis of the molybdenum cofactor. *J Biol Chem* **276**(37): 34695-701.

Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P. and Burge, C. B. (2003). Prediction of mammalian microRNA targets. *Cell* **115**(7): 787-798.

Lewis, J. E., Jensen, P. J. and Wheelock, M. J. (1994). Cadherin function is required for human keratinocytes to assemble desmosomes and stratify in response to calcium. *J Invest Dermatol* **102**(6): 870-7.

Lewis, J. E., Wahl 3rd, J. K., Sass, K. M., Jensen. P. J., Johnson, K. R. and Wheelock, M. J. (1997). Cross-talk between adherens junctions and desmosomes depends on plakoglobin. *J Cell Biol* **136**(4): 919-34.

Li, W., Bengtson, M. H., Ulbrich, A., Matsuda, A., Reddy, V. A., Orth, A., Chanda, S. K., Batalov, S. and Joazeiro, C. A. P. (2008). Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling. *PLoS One* **3**(1): e1487.

Lickert, H., Bauer, A., Kemler, R. and Stappert, J. (2000). Casein kinase II phosphorylation of E-cadherin increases E-cadherin/ß-catenin interaction and strengthens cell-cell adhesion. *J Biol Chem* **275**(7): 5090-5.

Lilien, J. and Balsamo, J. (2005). The regulation of cadherin-mediated adhesion by tyrosine phosphorylation/dephosphorylation of ß-catenin. *Curr Opin Cell Biol* **17**(5): 459-65.

Lim, L. P., Lau, N. C., Weinstein, E. G., Abdelhakim, A., Yekta, S., Rhoades, M. W., Burge, C. B. and Bartel, D. P. (2003). The mircroRNAs of Caenorhabditis elagans. *Genes Dev* **17**(8), 991-1008.

Liu, C., Li, Y., Semenov, M., Han, C., Baeg, G. H., Tan, Y., Zhang, Z., Lin, X. and He, X. (2002). Control of ß-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell* **108**(6): 837-47.

Liu, J., Stevens, J., Rote, C. A., Yost, H. J., Hu, Y., Neufeld, K. L., White, R. L. and Matsunami, N. (2001). Siah-1 mediates a novel β-catenin degradation pathway linking p53 to the adenomatous polyposis coli protein. *Mol Cell* **7**(5): 927-36.

Luhtala, N. and Odorizzi, G. (2004). Bro1 coordinates deubiquitination in the multivesicular body pathway by recruiting Doa4 to endosomes. *J Cell Biol* **166**(5): 717-29.

Lui, T. T., Lacroix, C., Ahmed, S. M., Goldenberg, S. J., Leach, C. A., Daulat, A. M., and Angers, S. (2011). The ubiquitin-specific protease USP34 regulates axin stability and Wnt/β-catenin signaling. *Mol Cell Biol* **31**(10): 2053-65.

MacDonald, B. T., Tamai, K., He, X. (2009). Wnt/ß-catenin signaling: components, mechanisms, and diseases. *Dev Cell* **17**(1): 9-26.

Machida, Y. J., Machida, Y., Vashisht, A. A., Wohlschlegel, J. A. and Dutta, A. (2009). The deubiquitinating enzyme BAP1 regulates cell growth via interaction with HCF-1. *J Biol Chem* **284**(49): 34179-88.

Mao, Y., Senic-Matuglia, F., Di Fiore, P. P., Polo, S., Hodsdon, M. E. and De Camilli, P. (2005). Deubiquitinating function of ataxin-3: insights from the solution structure of the Josephin domain. *Proc Natl Acad Sci U S A* **102**(36): 12700-5.

Marmor, M. D. and Yarden, Y. (2004). Role of protein ubiquitylation in regulating endocytosis of receptor tyrosine kinases. *Oncogene* **23**(11): 2057-70.

Martin, G. S. (2003). Cell signaling and cancer. *Cancer Cell* **4**(3): 167-74.

Matsumoto, M. L., Wickliffe, K. E., Dong, K. C., Yu, C., Bosanac, I., Bustos, D., Phu, L., Kirkpatrick, D. S., Hymowitz, S. G., Rape, M., Kelley, R. F. and Dixit, V. M. (2010). K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody. *Mol Cell* **39**(3): 477-84.

Matsuzawa, S. I. and Reed, J. C. (2001). Siah-1, SIP, and Ebi collaborate in a novel pathway for β-catenin degradation linked to p53 responses. *Mol Cell* **7**(5): 915-26.

Matteucci, E., Ridolfi, E. and Desiderio, M. A. (2006) Hepatocyte growth factor differently influences Met-E-cadherin phosphorylation and downstream signaling pathway in two models of breast cancer cells. *Cell. Mol. Life Sci.* **63**: 2016-2026.

McCrea, P. D., Turck, C. W. and Gumbiner, B. (1991). A homolog of the armadillo protein in Drosophila (plakoglobin) associated with E-cadherin. *Science* **254**(5036): 1359-61.

McCullough, J., Clague, M. J. and Urbé, S. (2004). AMSH is an endosome-associated ubiquitin isopeptidase. *J Cell Biol* **166**(4): 487-92.

McCullough, J., Row, P. E., Lorenzo, O., Doherty, M., Beynon, R., Clague, M. J. and Urbé, S. (2006). Activation of the endosome-associated ubiquitin isopeptidase AMSH by STAM, a component of the multivesicular body-sorting machinery. *Curr Biol* **16**(2): 160-5.

Messick, T. E., Russell, N. S., Iwata, A. J., Sarachan, K. L., Shiekhattar, R., Shanks, J. R., Reyes-Turcu, F. E., Wilkinson, K. D. and Marmorstein, R. (2008). Structural basis for ubiquitin recognition by the Otu1 ovarian tumor domain protein. *J Biol Chem* **283**(16): 11038-49.

Metzelaar, M. J., Wijngaard, P. L., Peters, P. J., Sixma, J. J., Nieuwenhuis, H. K. and Clevers, H. C. (1991) CD63 antigen. A novel lysosomal membrane glycoprotein, cloned by a screening procedure for intracellular antigens in eukaryotic cells. *J Biol Chem* **266:** 3239−3245

Metzger, M. B., Hristova, V. A. and Weissman, A. M. (2012). HECT and RING finger families of E3 ubiquitin ligases at a glance. *J Cell Sci* **125**, 531-537.

Miettinen, P. J., Ebner, R., Lopez, A. R. and Derynck, R. (1994). TGF-ß induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. *J Cell Biol* **127**(6 Pt 2): 2021-36.

Milovanovic, T., Planutis, K., Nguyen, A., Marsh, J. L., Lin, F., Hope, C. and Holcombe, R. F. (2004). Expression of Wnt genes and frizzled 1 and 2 receptors in normal breast epithelium and infiltrating breast carcinoma. *Int J Oncol* **25**(5): 1337-42.

Miranda, K. C., Joseph, S. R., Yap, A. S., Teasdale, R. D. and Stow, J. L. (2003). Contextual binding of p120ctn to E-cadherin at the basolateral plasma membrane in polarized epithelia. *J Biol Chem* **278**(44): 43480-8.

Miranda, K. C., Khromykh, T., Christy, P., Le, T. L., Gottardi, C. J., Yap, A. S., Stow, J. L. and Teasdale, R. D. (2001). A dileucine motif targets E-cadherin to the basolateral cell surface in Madin-Darby canine kidney and LLC-PK1 epithelial cells. *J Biol Chem* **276**(25): 22565-72.

Miura, H., Nishimura, K., Tsujimura, A., Matsumiya, K., Matsumoto, K., Nakamura, T. and Okuyama, A. (2001). Effects of hepatocyte growth factor on E-cadherin-mediated cell-cell adhesion in DU145 prostate cancer cells. *Urology* **58**(6): 1064-9.

Miyake, S., Lupher, M. L., Jr., Druker, B. and Band, H. (1998). The tyrosine kinase regulator Cbl enhances the ubiquitination and degradation of the platelet-derived growth factor receptor α. *Proc Natl Acad Sci U S A* **95**(14): 7927-32.

Mizuno, E., Kobayashi, K., Yamamoto, A., Kitamura, N. and Komada, M. (2006). A deubiquitinating enzyme UBPY regulates the level of protein ubiquitination on endosomes. *Traffic* **7**(8): 1017-31.

Morin, P. J. (1999). β-catenin signaling and cancer. *Bioessays* **21**(12): 1021-30.

Morin, P. J., Sparks, A. B., Korinek, V., Barker, N., Clevers, H., Vogelstein, B. and Kinzler, K. W. (1997). Activation of β-catenin-Tcf signaling in colon cancer by mutations in β-catenin or APC. *Science* **275**(5307): 1787-90.

Morris, J. R. and Solomon, E. (2004). BRCA1 : BARD1 induces the formation of conjugated ubiquitin structures, dependent on K6 of ubiquitin, in cells during DNA replication and repair. Hum *Mol Genet* **13**(8): 807-17.

Mosimann, C., Hausmann, G. and Basler, K. (2009). ß-catenin hits chromatin: regulation of Wnt target gene activation. *Nat Rev Mol Cell Biol* **10**(4): 276-86.

Mu, F., Callaghan, J. M., Steele-Mortimer, O., Stenmark, H., Parton, R. G., Campbell, P. L., McCluskey, J., Yeo, J., Tock, E. P. C. and Toh, B. (1995) EEA1, an early endosome-associated protein. *Journal of Biochmistry* **270**: 13503 - 13511.

Mueller, T. D. and Feigon, J. (2002). Solution structures of UBA domains reveal a conserved hydrophobic surface for protein-protein interactions. *J Mol Biol* **319**(5): 1243-55.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986) Specific enzymatic amplification of DNA in vitro: the Polymerase Chain Reaction. *Cold Spring Harbor Symposia on Quantitative Biology* **51**(1): 263-273.

Muratani, M. and Tansey, W. P. (2003). How the ubiquitin-proteasome system controls transcription. *Nat Rev Mol Cell Biol* **4**(3): 192-201.

Muroi, M., Shiragami, N., Nagao, K., Yamasaki, M., and Takatsuki, A. (1993). Folimycin (concanamycin A), a specific inhibitor of V-ATPase, blocks intracellular translocation of the glycoprotein of vesicular stomatitis virus before arrival to the Golgi apparatus. *Cell Struct Funct* **18**(3): 139-49.

Murray, R. Z., Jolly, L. A. and Wood, S. A. (2004) The FAM deubiquitylating enzyme localises to multiple points of protein trafficking in epithelia, where it associates with E-cadherin and β-catenin. *Mol Biol Cell* **15**: 1591-1599.

Nakamura, N. and Hirose, S. (2008). Regulation of Mitochondrial Morphology by USP 30, a Deubiquitinating Enzyme Present in the Mitochondrial Outer Membrance. *Mol Biol Cell* **19**(5): 1903-1911.

Nakamura, M., Tanaka, N., Kitamura, N. and Komada, M. (2006). Clathrin anchors deubiquitinating enzymes, AMSH and AMSH-like protein, on early endosomes. *Genes Cells* **11**(6): 593-606.

Nicastro, G., Masino, L., Esposito, V., Menon, R. P., De Simone, A., Fraternali, F. and Pastore, A. (2009). Josephin domain of ataxin-3 contains two distinct ubiquitin-binding sites. *Biopolymers* **91**(12): 1203-14.

Nicastro, G., Menon, R. P., Masino, L., Knowles, P. P., McDonald, N. Q. and Pastore, A. (2005). The solution structure of the Josephin domain of ataxin-3: structural determinants for molecular recognition. *Proc Natl Acad Sci U S A* **102**(30): 10493-8.

Pillay, C. S., Elliott, E. and Dennison, C. (2002). Endolysosomal proteolysis and its regulation. *Biochem J* **363**(Pt 3): 417-29.

Niessen, C. M. (2007). Tight junctions/adherens junctions: basic structure and function. *J Invest Dermatol* **127**(11): 2525-32.

Niessen, C. M. and Gottardi, C. J. (2008). Molecular components of the adherens junction. *Biochim Biophys Acta* **1778**(3): 562-71.

Nishikawa, H., Wu, W., Koike, A., Kojimo, R., Gomi, H., Fukuda, M. and Ohta, T. (2009). BRCA1-associated protein 1 interferes with BRCA1/BARD1 RING heterodimer activity. *Cancer Res* **69**(1): 111-9.

Nishisho, I., Nakamura, Y., Miyoshi, Y., Miki, Y., Ando, H., Horii, A., Koyama, K., Utsunomiya, J., Baba, S. and Hedge, P. (1991). Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science* **253**(5020): 665-9.

Noordermeer, J., Klingensmith, J., Perrimon, N. and Nusse, R. (1994). dishevelled and armadillo act in the wingless signalling pathway in Drosophila. *Nature* **367**(6458): 80-3.

Nose, A., Nagafuchi, A., and Takeichi, M. (1988). Expressed recombinant cadherins mediate cell sorting in model systems. *Cell* **54**(7): 993-1001.

Nose, A., Tsuji, K. and Takeichi, M. (1990). Localization of specificity determining sites in cadherin cell adhesion molecules. *Cell* **61**(1): 147-55.

Nüsslein-Volhard, C. and Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. *Nature* **287**(5785): 795-801.

Paddison P. J., Silva, J. M., Conklin, D. S., Schlabach, M., Li, M., Aruleba, S., Balija, V., O'Shaughnessy, A., Gnoj, L., Scobie, K., Chang, K., Westbrook, T., Cleary, M., Sachidanandam, R., McCombie, W. R., Elledge, S. J. and Hannon, G. J. (2004). A resource for large-scale RNA-interference-based screens in mammals. *Nature* **428**: 427–31.

Poole, B., Ohkuma, S. and Warburton, M. J. (1977). The accumulation of weakly basic substances in lysosomes and the inhibition of intracellular protein degradation. *Acta Biol Med Ger* **36**(11-12): 1777-88.

Oda, T., Kanai, Y., Oyama, T., Yoshiura, K., Shimoyama, Y., Birchmeier, W., Sugimura, T. and Hirohashi, S. (1994). E-cadherin gene mutations in human gastric carcinoma cell lines. *Proc Natl Acad Sci U S A* **91**(5): 1858-62.

Oda, T., Kanai, Y., Shimoyama, Y., Nagafuchi, A., Tsukita, S. and Hirohashi, S. (1993). Cloning of the human α-catenin cDNA and its aberrant mRNA in a human cancer cell line. *Biochem Biophys Res Commun* **193**(3): 897-904.

Ohkubo, T. and Ozawa, M. (1999). p120(ctn) binds to the membrane-proximal region of the E-cadherin cytoplasmic domain and is involved in modulation of adhesion activity. *J Biol Chem* **274**(30): 21409-15.

Ohshima, R., Ohta, T., Wu, W., Koike, A., Iwatani, T., Henderson, M., Watts, C. K. and Otsubo, T. (2007). Putative tumor suppressor EDD interacts with and up-regulates APC. *Genes Cells* **12**(12): 1339-45.

Opipari, A. W., Jr., Boguski, M. S. and Dixit, V. M. (1990). The A20 cDNA induced by tumor necrosis factor alpha encodes a novel type of zinc finger protein. *J Biol Chem* **265**(25): 14705-8.

Orford, K., Crockett, C., Jensen, J. P., Weissman, A. M. and Byers, S. W. (1997). Serine phosphorylation-regulated ubiquitination and degradation of β-catenin. *J Biol Chem* **272**(40): 24735-8.

Ott, D. E., Coren, L. V., Copeland, T. D., Kane, B. P., Johnson, D. G., Sowder, R. C., 2nd, Yoshinaka, Y., Oroszlan, S., Arthur, L. O. and Henderson, L. E. (1998). Ubiquitin is covalently attached to the p6Gag proteins of human immunodeficiency virus type 1 and simian immunodeficiency virus and to the p12Gag protein of Moloney murine leukemia virus. *J Virol* **72**(4): 2962-8.

Oyama, T., Kanai, Y., Ochiai, A., Akimoto, S., Oda, T., Yanagihara, K., Nagafuchi, A., Tsukita, S., Shibamoto, S., Ito, F., Takeichi, M., Matsuda, H. and Hirohashi, S. (1994). A truncated β-catenin disrupts the interaction between E-cadherin and α-catenin: a cause of loss of intercellular adhesiveness in human cancer cell lines. *Cancer Res* **54**(23): 6282-7.

Ozawa, M., Baribault, H. and Kemler, R. (1989). The cytoplasmic domain of the cell adhesion molecule uvomorulin associates with three independent proteins structurally related in different species. *EMBO J* **8**(6): 1711-7.

Ozawa, M., Hoschützky, H., Herrenknecht, K. and Kemler, R. (1990). A possible new adhesive site in the cell-adhesion molecule uvomorulin. *Mech Dev* **33**(1): 49-56.

Ozawa, M., Ringwald, M. and Kemler, R. (1990). Uvomorulin-catenin complex formation is regulated by a specific domain in the cytoplasmic region of the cell adhesion molecule. *Proc Natl Acad Sci U S A* **87**(11): 4246-50.

Palombella, V. J., O. J. Rando, et al. (1994). The ubiquitin-proteasome pathway is required for processing the NF-kappa B1 precursor protein and the activation of NF-κ B. *Cell* **78**(5): 773-85.

Paterson, A. D., Parton, R. G., Ferguson, C., Stow, J. L. and Yap, A. S. (2003). Characterization of E-cadherin endocytosis in isolated MCF-7 and chinese hamster ovary cells: the initial fate of unbound E-cadherin. *J Biol Chem* **278**(23): 21050-7.

Peifer, M., Pai, L. M. and Casey, M. (1994). Phosphorylation of the Drosophila adherens junction protein Armadillo: roles for wingless signal and zeste-white 3 kinase. *Dev Biol* **166**(2): 543-56.

Peifer, M., Sweeton, D., Casey, M. and Wieschaus, E. (1994). wingless signal and Zeste-white 3 kinase trigger opposing changes in the intracellular distribution of Armadillo. *Development* **120**(2): 369-80.

Peifer, M. and Yap, A. S. (2003). Traffic control: p120-catenin acts as a gatekeeper to control the fate of classical cadherins in mammalian cells. *J Cell Biol* **163**(3): 437-40.

Peña-Llopis, S., Vega-Rubin-de-Celis, S., Liao, A., Leng, N., Pavía-Jiménez, A., Wang, S., Yamasaki, T., Zhrebker, L., Sivanand, S., Spence, P., Kinch, L., Hambuch, T., Jain, S., Lotan, Y., Margulis, V., Sagalowsky, A. I., Summerour, P. B., Kabbani, W., Wong, S. W. W., Grishin, N., Laurent, M., Xie, X., Haudenschild, C. D., Ross, M. T., Bentley, D. R., Kapur, P. and Brugarolas, J. (2012). BAP1 loss defines a new class of renal cell carcinoma. *Nat Genet* **44**(7): 751-9.

Peng, J., Schwartz, D., Elias, J. E., Thoreen, C. C., Cheng, D., Marsischky, G., Roelofs, J., Finley, D. and Gygi, S. P. (2003). A proteomics approach to understanding protein ubiquitination. *Nat Biotechnol* **21**(8): 921-6.

Pérez-Moreno, M. A., Locascio, A., Rodrigo, I., Dhondt, G., Portillo, F., Nieto, M. A. and Cano, A. (2001). A new role for E12/E47 in the repression of E-cadherin expression and epithelial-mesenchymal transitions. *J Biol Chem* **276**(29): 27424-31.

Perl, A. K., Wilgenbus, P., Dahl, U., Semb, H. and Christofori, G. (1998). A causal role for E-cadherin in the transition from adenoma to carcinoma. *Nature* **392**(6672): 190-3.

Pettitt, J. (2005). The cadherin superfamily. *WormBook*: 1-9.

Piedra, J., Miravet, S., Castaño, J., Pálmer, H. G., Heisterkamp, N., de Herreros, A. G. and Duñach, M. (2003). p120 Catenin-associated Fer and Fyn tyrosine kinases regulate β-catenin Tyr-142 phosphorylation and β-catenin-α-catenin Interaction. *Mol Cell Biol* **23**(7): 2287-97.

Pickart, C. M. and Cohen, R. E. (2004). Proteasomes and their kin: proteases in the machine age. *Nat Rev Mol Cell Biol* **5**(3): 177-87.

Pokutta, S. and Weis, W. I. (2000). Structure of the dimerization and β-catenin-binding region of α-catenin. *Mol Cell* **5**(3): 533-43.

Polakis, P. (2000). Wnt signaling and cancer. *Genes Dev* **14**(15): 1837-51.

Qiu, X. B., Ouyang, S. Y., Li, C. J., Miao, S., Wang, L. and Goldberg, A. L. (2006). hRpn13/ADRM1/GP110 is a novel proteasome subunit that binds the deubiquitinating enzyme, UCH37. *EMBO J* **25**(24): 5742-53.

Raasi, S., Varadan, R., Fushman, D. and Pickart, C. M. (2005). Diverse polyubiquitin interaction properties of ubiquitin-associated domains. *Nat Struct Mol Biol* **12**(8): 708-14.

Rabinovitz, M. and Fisher, J. M. (1964). Characteristics of the Inhibition of Hemoglobin Synthesis in Rabbit Reticulocytes by Threo-α-Amino-β-Chlorobutyric Acid. *Biochim Biophys Acta* **91**: 313-22.

Rahighi, S., Ikeda, F., Kawasaki, M., Akutsu, M., Suzuki, N., Kato, R., Kensche, T., Uejima, T., Bloor, S., Komander, D., Randow, F., Wakatsuki, S. and Dikic, I. (2009). Specific recognition of linear ubiquitin chains by NEMO is important for NF-kappaB activation. *Cell* **136**(6): 1098-109.

Raiborg, C. and Stenmark, H. (2009) The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins. *Nature* **458**: 445 - 452.

Raiborg, C. and Stenmark, H. (2009). The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins. *Nature* **458**(7237): 445-52.

Riess, O., Rüb, U., Pastore, A., Bauer, P. and Schöls L. (2008). SCA3: neurological features, pathogenesis and animal models. *Cerebellum* **7**(2): 125-37.

Reshetnikova, G., Troyanovsky, S. and Rimm, D. L. (2007). Definition of a direct extracellular interaction between Met and E-cadherin. *Cell Biol Int* **31**(4): 366-73.

Reya, T. and Clevers, H. (2005). Wnt signalling in stem cells and cancer. *Nature* **434**(7035): 843-50.

Reynolds, A. B. and Carnahan, R. H. (2004). Regulation of cadherin stability and turnover by p120ctn: implications in disease and cancer. *Semin Cell Dev Biol* **15**(6): 657-63.

Reynolds, A. B., Daniel, J. M., Mo, Y., Wu, J. and Zhang, Z. (1996). The novel catenin p120cas binds classical cadherins and induces an unusual morphological phenotype in NIH3T3 fibroblasts. *Exp Cell Res* **225**(2): 328-37.

Rhee, C. S., Sen, M., Lu, D., Wu, C., Leoni, L., Rubin, J., Corr, M. and Carson, D. A. (2002). Wnt and frizzled receptors as potential targets for immunotherapy in head and neck squamous cell carcinomas. *Oncogene* **21**(43): 6598-605.

Rhee, J., Mahfooz, N. S., Arregui, C., Lilien, J., Balsamo, J. and VanBerkum, M. F. A. (2002). Activation of the repulsive receptor Roundabout inhibits N-cadherin-mediated cell adhesion. *Nat Cell Biol* **4**(10): 798-805.

Riggleman, B., Wieschaus, E. and Schedl, P. (1989). Molecular analysis of the armadillo locus: uniformly distributed transcripts and a protein with novel internal repeats are associated with a Drosophila segment polarity gene. *Genes Dev* **3**(1): 96-113.

Rimm, D. L., E. R. Koslov, et al. (1995). Alpha1(E)-catenin is an actin-binding and -bundling protein mediating the attachment of F-actin to the membrane adhesion complex. *Proc Natl Acad Sci USA* **92**(19): 8813-7.

Rimm, D. L. and Morrow, J. S. (1994). Molecular cloning of human E-cadherin suggests a novel subdivision of the cadherin superfamily. *Biochem Biophys Res Commun* **200**(3): 1754-61.

Rosivatz, E., Becker, I., Specht, K, Fricke, E., Luber, B, Busch, R., Höfler, H. and Becker, K. F. (2002). Differential expression of the epithelial-mesenchymal transition regulators snail, SIP1, and twist in gastric cancer. *Am J Pathol* **161**(5): 1881-91.

Rotin, D. and Kumar, S. (2009). Physiological functions of the HECT family of ubiquitin ligases. *Nat Rev Mol Cell Biol* **10**(6): 398-409.

Row, P. E., Prior, I. A., McCullough, J., Clague, M. J. and Urbé, S. (2006). The ubiquitin isopeptidase UBPY regulates endosomal ubiquitin dynamics and is essential for receptor down-regulation. *J Biol Chem* **281**(18): 12618-24.

Rubinfeld, B., Albert, I., Porfiri, E., Fiol, C., Munemitsu, S. and Polakis, P. (1996). Binding of GSK3β to the APC-β-catenin complex and regulation of complex assembly. *Science* **272**(5264): 1023-6.

Rubinfeld, B., Souza, B., Albert. I. Müller, O, Chamberlain, S. H., Masiarz, F. R., Menemitsu, S. and Polakis, P (1993). Association of the APC gene product with β-catenin. *Science* **262**(5140): 1731-4.

Rubinfeld, B., Souza, B., Albert, I., Munemitsu, S. and Polakis, P. (1995). The APC protein and E-cadherin form similar but independent complexes with α-catenin, β-catenin, and plakoglobin. *J Biol Chem* **270**(10): 5549-55.

Sacco, J. J., Coulson, J. M., Clague, M. J. and Urbé, S. (2010). Emerging roles of deubiquitinases in cancer-associated pathways. *IUBMB Life* **62**(2): 140-57.

Sakai, T., Toguchida, J., Ohtani, N., Yandell, D. W., Rapaport, J. M. and Dryja, T. P. (1991). Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. *Am J Hum Genet* **48**(5): 880-8.

Saitoh, M., Shirakihara, T. and Miyazono, K. (2009). Regulation of the stability of cell surface E-cadherin by the proteasome. *Biochem Biophys Res Commun* **381**(4): 560-5.

Sato, Y., Yoshikawa, A., Yamagata, A., Mimura, H., Yamashita, M., Ookata, K., Nureki, O., Iwai, K., Komada, M. and Fukai, S. (2008). Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains. *Nature* **455**, 358-362

Satoh, S., Daigo, Y., Furukawa, Y., Kato, T., Miwa, N, Nishiwaki, T., Kawasoe, T., Ishiguro, H., Fujita, M., Tokino, T., Sasaki, Y., Imaoka, S., Murata, M., Shimano, T., Yamaoka, Y. and Nakamura, Y. (2000). AXIN1 mutations in hepatocellular carcinomas, and growth suppression in cancer cells by virus-mediated transfer of AXIN1. *Nat Genet* **24**(3): 245-50.

Scheel, H., Tomiuk, S. and Hofmann, K. (2003). Elucidation of ataxin-3 and ataxin-7 function by integrative bioinformatics. *Hum Mol Genet* **12**(21): 2845-52.

Schimke, R. T. and Doyle, D. (1970). Control of enzyme levels in animal tissues. *Annu Rev Biochem* **39**: 929-76.

Schoenheimer, R., Ratner, S. and Rittenberg, D. (1939). The Process of Continuous Deamination and Reamination of Amino Acids in the Proteins of Normal Animals. *Science* **89**(2308): 272-3.

Schulman, B. A. and Harper, J. W. (2009). Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat Rev Mol Cell Biol* **10**(5): 319-31.

Schwarz-Romond, T., Fiedler, M., Shibata, N., Butler, P. J., Kikuchi, A., Higuchi, Y. and Bienz, M. (2007). The DIX domain of Dishevelled confers Wnt signaling by dynamic polymerization. *Nat Struct Mol Biol* **14**(6): 484-92.

Semizarov, D., Frost, L., Sarthy, A., Kroeger, P., Halbert, D. N. and Fesil, S. W. (2003). Specificity of short interfering RNA determined through gene expression signature. *Proc Natl Acad Sci USA* **100**(11): 6347-6352.

Seshadri, T., Uzman, J. A, Oshima, J. and Campisi, J. (1993) Identification of a transcript that is down-regulated in senescent human fibroblasts. Cloning, sequence analysis, and regulation of the human L7 ribosomal protein gene. *J Biol Chem* **268**: 18474–80.

Shimoyama, Y., Nagafuchi, A., Fujita, S., Gotoh, M., Takeichi, M., Tsukita, S. and Hiroshi, S. (1992). Cadherin dysfunction in a human cancer cell line: possible involvement of loss of α-catenin expression in reduced cell-cell adhesiveness. *Cancer Res* **52**(20): 5770-4.

Siegfried, E., Chou, T. B. and Perrimon, N. (1992). wingless signaling acts through zeste-white 3, the Drosophila homolog of glycogen synthase kinase-3, to regulate engrailed and establish cell fate. *Cell* **71**(7): 1167-79.

Sierra, J., Yoshida, T., Joazeiro, C. A. and Jones, K. A. (2006). The APC tumor suppressor counteracts β-catenin activation and H3K4 methylation at Wnt target genes. *Genes Dev* **20**(5): 586-600.

Siomi, H. and Siomi, M. C. (2009). On the road to reading the RNA-interference code. *Nature* **457**(7228): 396-404.

Sioud, M. (2005). Introduction of inflammatory cytokines and interferon respnses by double-stranded and single-stranded siRNAs is sequence-dependent and requires endosomal localisation. *J Mol Biol* **348**(5): 1079-1090.

Skaug, B., Jiang, X., Chen and Z. J. (2009). The role of ubiquitin in NF-κB regulatory pathways. *Annu Rev Biochem* **78**: 769-96.

Sobhian, B., Shao, G., Lilli, D. R., Culhane, A. C., Moreau, L. A., Xia, B., Livingston, D. M and Greenberg, R. A. (2007). RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* **316**(5828): 1198-202.

Song, E. J., Werner, S. L., Neubauer, J., Stegmeier, F., Aspden, J., Rio, D., Harper, J. W., Elledge, S. J., Kirschner, M. W. and Rape, M. (2010). The Prp19 complex and the Usp4Sart3 deubiquitinating enzyme control reversible ubiquitination at the spliceosome. *Genes Dev* **24**(13): 1434-47.

Song, H. Y., Rothe, M. and Goeddel, D. V. (1996). The tumor necrosis factor-inducible zinc finger protein A20 interacts with TRAF1/TRAF2 and inhibits NF-kappaB activation. *Proc Natl Acad Sci U S A* **93**(13): 6721-5.

Sowa, M. E., Bennet, E. J., Gygi, S. P. and Harper, J. W. (2009) Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**: 389 – 403.

Sparks, A. B., Morin, P. J., Vogelstein, B. and KInzler, K. W. (1998). Mutational analysis of the APC/β-catenin/Tcf pathway in colorectal cancer. *Cancer Res* **58**(6): 1130-4.

St-Pierre, Y., Campion, C. G. and Grosset, A. (2012). A distinctive role of galectine-7 in cancer? *Frontiers in Bioscience* **17**; 438-450.

Stang, E., Blystad, F. D., Kazazic, M., Bertelsen, V., Brodahl, T., Raiborg, C., Stenmark, H. and Madshus, I. H. (2004). Cbl-dependent ubiquitination is required for progression of EGF receptors into clathrin-coated pits. *Mol Biol Cell* **15**(8): 3591-604.

Stella, M. C. and P. M. Comoglio (1999). HGF: a multifunctional growth factor controlling cell scattering. *Int J Biochem Cell Biol* **31**(12): 1357-62.

Stelter, P. and Ulrich, H. D. (2003). Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. *Nature* **425**(6954): 188-91.

Stoker, M. and Perryman, M. (1985). An epithelial scatter factor released by embryo fibroblasts. *J Cell Sci* **77**: 209-23.

Storer, A. C. and Menard, R. (1994). Catalytic mechanism in papain family of cysteine peptidases. *Methods Enzymol* **244**: 486-500.

Strous, G. J., van Kerkhof, P., Govers, R., Ciechanover, A. and Schwartz, A. L. (1996). The ubiquitin conjugation system is required for ligand-induced endocytosis and degradation of the growth hormone receptor. *EMBO J* **15**(15): 3806-12.

Su, L. K., Vogelstein, B. and Kinzler, K. W. (1993). Association of the APC tumor suppressor protein with catenins. *Science* **262**(5140): 1734-7.

Suzuki, H., Watkins, D. N., Jair, K. W., Schuebel, K. E., Markowitz, S. D., Chen, W. D., Pretlow, T. P., Yang, B., Akiyama, Y., van Engeland, M., Toyota, M., Tokino, T., Hinoda, Y., Imai, K., Herman, J. G. and Baylin, S. B. (2004). Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer. *Nat Genet* **36**(4): 417-22.

Swaminathan, G. and C. A. Cartwright (2012). Rack1 promotes epithelial cell-cell adhesion by regulating E-cadherin endocytosis. *Oncogene* **31**(3): 376-89.

Takeichi, M. (1977). Functional correlation between cell adhesive properties and some cell surface proteins. *J Cell Biol* **75**(2 Pt 1): 464-74.

Swanson, K. A., Kang, R. S., Stamenova, S. D., Hicke, L. and Radhakrishnan, I. (2003). Solution structure of Vps27 UIM-ubiquitin complex important for endosomal sorting and receptor downregulation. *EMBO J* **22**(18): 4597-606.

Tanaka, K., Waxman, L. and Goldberg, A. L. (1983). ATP serves two distinct roles in protein degradation in reticulocytes, one requiring and one independent of ubiquitin. *J Cell Biol* **96**(6): 1580-5.

Tauriello, D. V., Haegebarth, A., Kuper, I., Edelmann, M. J., Henraat, M., Canninga-van Dijk, M. R., Kessler, B. M., Clevers, H. and Maurice, M. M. (2010). Loss of the tumor suppressor CYLD enhances Wnt/β-catenin signaling through K63-linked ubiquitination of Dvl. *Mol Cell* **37**(5): 607-19.

Tauriello, D. V. and Maurice, M. M. (2010). The various roles of ubiquitin in Wnt pathway regulation. *Cell Cycle* **9**(18): 3700-9.

Taya, S., Yamamoto, T., Kano, K., Kawano, Y., Iwamatsu, A., Tsuchiya, T., Tanaka, K., Kanai-Azuma, M., Wood, S. A., Mattick, J. S. and Kaibuchi, K. (1998) The Ras target AF-6 is a substrate of the fam deubiquitinating enzyme. *J Cell Biol* **142**(4): 1053 – 1062.

Taya, S., Yamamoto, T., Kanai-Azuma, M., Wood, S. A. and Kaibuchi, K. (1999) The deubiquitinating enzyme FAM interacts with and stabilises b-catenin. *Genes Cells* **4**: 757-767.

Taylor, S. V., Kelleher, N. L., Kinsland, C., Chiu, H. J., Costello, C. A., Backstorm, A. D., McLafferty, F. W. and Begley, T. P. (1998). Thiamin biosynthesis in Escherichia coli. Identification of ThiS thiocarboxylate as the immediate sulfur donor in the thiazole formation. *J Biol Chem* **273**(26): 16555-60.

Terrell, J., Shih, S., Dunn, R. and Hicke, L. (1998). A function for monoubiquitination in the internalization of a G protein-coupled receptor. *Mol Cell* **1**(2): 193-202.

Testa, J. R., Cheung, M., Pei, J., Below, J. E., Tan, Y., Sementino, E., Cox, N. J., Dogan, A. U., Pass, H. I., Trusa, S., Hesdorffer, M., Nasu, M., Powers, A., Rivera, Z., Comertpay, S., Tanji, M., Gaudino, G., Yang, H. and Carbone, M. (2011). Germline BAP1 mutations predispose to malignant mesothelioma. *Nat Genet* **43**(10): 1022-5.

Thiery, J. P. (2003). Epithelial-mesenchymal transitions in development and pathologies. *Curr Opin Cell Biol* **15**(6): 740-6.

Thiery, J. P. and Sleeman, J. P. (2006). Complex networks orchestrate epithelial-mesenchymal transitions. *Nat Rev Mol Cell Biol* **7**(2): 131-42.

Thoreson, M. A. and Reynolds, A. B. (2002). Altered expression of the catenin p120 in human cancer: implications for tumor progression. *Differentiation* **70**(9-10): 583-9.

Thorne, C., Eccles, R. L., Coulson, J. M., Urbé, S. and Clague, M. J. (2011). Isoform-Specific Localisation of the Deubiquitinase USP33 to the Golgi Apparatus. *Traffic* **12**(11): 1563-1574

Todi, S. V., Winborn, B. J., Scaglione, K. M., Blount, J. R., Travis, S. M. and Paulson, H. L. (2009). Ubiquitination directly enhances activity of the deubiquitinating enzyme ataxin-3. *EMBO J* **28**(4): 372-82.

Tokunaga, F., Nakagawa, T., Nakahara, M., Saeki, Y., Taniguchi, M., Sakata, S., Tanaka, K., Nakano, H. and Iwai, K. (2011). SHARPIN is a component of the NF-κB-activating linear ubiquitin chain assembly complex. *Nature* **471**(7340): 633-6.

Tran, H., Hamada, F., Schwarz-Romond, T and Bienz, M. (2008). Trabid, a new positive regulator of Wnt-induced transcription with preference for binding and cleaving K63-linked ubiquitin chains. *Genes Dev* **22**(4): 528-42.

Tunggal, J. A., Helfrich, I., Schmitz, A., Schwarz, H., Günzel, D., Fromm, M., Kemler, R., Krieg, T., and Niessen, C. M. (2005). E-cadherin is essential for in vivo epidermal barrier function by regulating tight junctions. *EMBO J* **24**(6): 1146-56.

Tuschl, T. Zamore, P. D., Lehmann, R., Bartel, D. P. and Sharp, P. A. (1999). Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev* **13**, 3191-3197.

Tyagi, S., Chabes, A. L., Wysocka, J. and Herr, W. (2007). E2F activation of S phase promoters via association with HCF-1 and the MLL family of histone H3K4 methyltransferases. *Mol Cell* **27**(1): 107-19.

Uematsu, K., He, B., You, L., Xu, Z., McCormick, F. and Jablons, D. M. (2003). Activation of the Wnt pathway in non small cell lung cancer: evidence of dishevelled overexpression. *Oncogene* **22**(46): 7218-21.

Uematsu, K., Kanazawa, S., You, L., He, B., Xu, Z., Li, K., Peterlin, B. M., McCormick, F. and Jablons, D. M. (2003). Wnt pathway activation in mesothelioma: evidence of Dishevelled overexpression and transcriptional activity of β-catenin. *Cancer Res* **63**(15): 4547-51.

Ugolini, F., Charafe-Jauffret, E., Bardou, V. J., Geneix, J., Adélaïde, J., Labat-Moleur, F., Penault-Llora, F., Longy, M., Jacquemier, J., Birnham, D. and Pébusque, M. J. (2001). WNT pathway and mammary carcinogenesis: loss of expression of candidate tumor suppressor gene SFRP1 in most invasive carcinomas except of the medullary type. *Oncogene* **20**(41): 5810-7.

Urbé, S., Liu, H., Hayes, S. D., Heride, C., Rigden, D. J. and Clague, M. J. (2012) Systematic survey of deubiquitinase localization identifies USP21 as a regulator of centrosome- and microtubule-associated functions. *Mol. Biol. Cell* **23**: 1095 – 1103.

van der Horst, A., de Vries-Smits, A. M., Brenkman, A. B., van Triest, M. H., van den Broek, N., Colland, F., Maurice, M. M. and Burgering, B. M. (2006). FOXO4 transcriptional activity is regulated by monoubiquitination and USP7/HAUSP. *Nat Cell Biol* **8**(10): 1064-73.

van Roy, F. and Berx, G. (2008). The cell-cell adhesion molecule E-cadherin. *Cell Mol Life Sci* **65**(23): 3756-88.

Ventii, K. H., Devi, N. S., Friedrich, K. L., Chernova, T. A., Tighiouart, M., Van Meir, E. G. and Wilkinson, K. D. (2008). BRCA1-associated protein-1 is a tumor suppressor that requires deubiquitinating activity and nuclear localization. *Cancer Res* **68**(17): 6953-62.

Virdee, S., Ye, Y., Nguyen, D. P., Komander, D. and Chin, J. W. (2010). Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase. *Nat Chem Biol* **6**(10): 750-757.

Vogelstein, B. and Kinzler, K. W. (1993). The multistep nature of cancer. *Trends Genet* **9**(4): 138-41.

Vogelstein, B. and Kinzler, K. W. (2004). Cancer genes and the pathways they control. *Nat Med* **10**(8): 789-99.

Voges, D., Zwickl, P. and Baumeister, W. (1999). The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu Rev Biochem* **68**: 1015-68.

Wada, K. and Kamitani, T. (2006) UnpEL/Usp4 is ubiquitinated by Ro52 and deubiqutinated by itself. *Biochem. and Biophy. Res. Commun.* **342**(1) : 253-258.

Wang, B., Matsuoka, S., Ballif, B. A., Zhang, D., Smogorzewska, A., Gygi, S. P. and Elledge, S. J. (2007). Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science* **316**(5828): 1194-8.

Wang, C., Xi, J., Begley, T. P. and Nicholson, L. K. (2001). Solution structure of ThiS and implications for the evolutionary roots of ubiquitin. *Nat Struct Biol* **8**(1): 47-51.

Wang, W., Va De Water, T. and Lufkin, T. (1998) Inner ear and maternal reproductive defects in mice lacking the Hmx3 homeobox gene. *Development* **125**: 621-634.

Weake, V. M. and Workman, J. L. (2008). Histone ubiquitination: triggering gene activity. *Mol Cell* **29**(6): 653-63.

Weeks, S. D., Grasty, K. C., Hernandez-Cuebas, L. and Loll. P. J. (2011). Crystal structure of a Josephin-ubiquitin complex: evolutionary restraints on ataxin-3 deubiquitinating activity. *J Biol Chem* **286**(6): 4555-65.

Wei, W., Li, M., Wang, J., Nie, F. and Li, L. (2012). The E3 Ubiquitin Ligase ITCH Negatively Regulates Canonical Wnt Signaling by Targeting Dishevelled Protein. *Mol Cell Biol* **32**(19): 3903-12.

Weinberg, R. A. (2007). *The biology of cancer*. New York, Garland Science.

Wiesner, T., Obenauf, A. C., Murali, R., Fried, I., Griewank, K. G., Ulz, P., Windpassinger, C., Wackernagel, W., Loy, S., Wolf, I., Viale, A., Lash, A. E., Pirun, M., Socci, N. D., Rütten, A., Palmedo, G., Abramson, D., Offit, K., Ott, A., Becker, J. C., Cerroni, L., Kutzner, H., Bastian, B. C. and Speicher, M. R. (2011). Germline mutations in BAP1 predispose to melanocytic tumors. *Nat Genet* **43**(10): 1018-21.

Wenzel, D. M., Stoll, K. E. and Klevit, R. E. (2011). E2s: structurally economical and functionally replete. *Biochem J* **433**(1): 31-42.

Wertz, I. E., O'Rourke, K. M., Zhou, H, Eby, M., Aravind, L, Seshagiri, S., Wu, P., Wiesmann, C., Baker, R., Boone, D. L., Ma. A., Koonin, E. V. and Dixit, V. M. (2004). De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF-κB signalling. *Nature* **430**(7000): 694-9.

White, A. E. and Harper, J. W. (2012) Emerging anatomy of the BAP1 tumour suppressor system. *Science* **337**: 1463-1464.

Wiborg, O., Pedersen, M. S., Wind, A, Berglund, L. E., Marcker, K. A. and Vuust, J. (1985). The human ubiquitin multigene family: some genes contain multiple directly repeated ubiquitin coding sequences. *EMBO J* **4**(3): 755-9.

Wightman, B., Ha, I. and Ruvkun, G. (1993). Posttranscriptional regulational of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell* **75**, 855-862.

Wilkinson, C. R., Seeger, M., Hartmann-Petersen, R., Stone, M., Wallace, M., Semple, C. and Gordon, C. (2001). Proteins containing the UBA domain are able to bind to multi-ubiquitin chains. *Nat Cell Biol* **3**(10): 939-43.

Wilkinson, K. D., Urban, M. K. and Hass, A. L. (1980). Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. *J Biol Chem* **255**(16): 7529-32.

Willert, K. and Nusse, R. (1998). β-catenin: a key mediator of Wnt signaling. *Curr Opin Genet Dev* **8**(1): 95-102.

Williams, R. L. and Urbe, S. (2007). The emerging shape of the ESCRT machinery. *Nat Rev Mol Cell Biol* **8**(5): 355-68.

Wong, E. and Cuervo, A. M. (2010). Integration of clearance mechanisms: the proteasome and autophagy. *Cold Spring Harb Perspect Biol* **2**(12): a006734.

Wu, R. C., Feng, Q., Lonard, D. M. and O'Malley, B. (2007). SRC-3 coactivator functional lifetime is regulated by a phospho-dependent ubiquitin time clock. *Cell* **129**(6): 1125-40.

Wu, T., Merbl, Y., Huo, Y., Gallop, J. L., Tzur, A. and Kirschner, M. W. (2010). UBE2S drives elongation of K11-linked ubiquitin chains by the anaphase-promoting complex. *Proc Natl Acad Sci U S A* **107**(4): 1355-60.

Xiao, K., Allison, D. F., Buckley, K. M., Kottke, M. D., Vincent, P. A., Faundez, V. and Kowalczyk, A. P. (2003). Cellular levels of p120 catenin function as a set point for cadherin expression levels in microvascular endothelial cells. *J Cell Biol* **163**(3): 535-45.

Xiao, K., Allison, D. F., Kottke, M. D., Summers, S., Sorescu, G. P., Faundez, V. and Kowalczyk, A. P. (2003). Mechanisms of VE-cadherin processing and degradation in microvascular endothelial cells. *J Biol Chem* **278**(21): 19199-208.

Xiao, K., Oas, R. G., Chiasson, C. M. and Kowalczyk, A. P. (2007). Role of p120-catenin in cadherin trafficking. *Biochim Biophys Acta* **1773**(1): 8-16.

Xiao, N., Li, H., Luo, J., Wang, R., Chen, H., Chen, J. and Wang, P. (2012). Ubiquitin-specific protease 4 (USP4) targets TRAF2 and TRAF6 for deubiquitination and inhibits TNFalpha-induced cancer cell migration. *Biochem J* **441**(3): 979-86.

Xing, Y., Clements, W. K., Kimelman, D. and Xu, W. (2003). Crystal structure of a β-catenin/axin complex suggests a mechanism for the β-catenin destruction complex. *Genes Dev* **17**(22): 2753-64.

Xu, P., Duong, D. M., Seyfried, N. T., Cheng, D., Xie, Y., Robert, J., Rush, J., Hochstrasser, M., Finley, D. and Peng, J. (2009). Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. *Cell* **137**(1): 133-45.

Yamada, S., Pokutta, S., Drees, F., Weis, W. I. and Nelson, W. J. (2005). Deconstructing the cadherin-catenin-actin complex. *Cell* **123**(5): 889-901.

Yang, J., Mani, S. A., Donaher, J. L., Ramaswamy, S., Itzykson, R. A., Come, C., Savagner, P., Gitelman, I., Richardson, A. and Weinberg, R. A. (2004). Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* **117**(7): 927-39.

Yang, J. and Weinberg, R. A. (2008). Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev Cell* **14**(6): 818-29.

Yao, T., Song, L., Xu, W., DeMartino, G. N., Florens, L., Swanson, S. K., Washburn, M. P., Conaway, J. W. and Cohen, R. E. (2006). Proteasome recruitment and activation of the Uch37 deubiquitinating enzyme by Adrm1. *Nat Cell Biol* **8**(9): 994-1002.

Yap, A. S., Brieher, W. M. and Gumbiner, B. M. (1997). Molecular and functional analysis of cadherin-based adherens junctions. *Annu Rev Cell Dev Biol* **13**: 119-46.

Yap, A. S., Crampton, M. S. and Hardin, J. (2007). Making and breaking contacts: the cellular biology of cadherin regulation. *Curr Opin Cell Biol* **19**(5): 508-14.

Ye, Y., Scheel, H., Hofmann, K. and Komander, D. (2009) Dissection of USP catalytic domain reveals five common insertion points. *Molecular BioSystems* **5**: 1797-1808.

Ye, Y. and Rape, M. (2009). Building ubiquitin chains: E2 enzymes at work. *Nat Rev Mol Cell Biol* **10**(11): 755-64.

Yoshida-Noro, C., Suzuki, N. and Takeichi, M. (1984). Molecular nature of the calcium-dependent cell-cell adhesion system in mouse teratocarcinoma and embryonic cells studied with a monoclonal antibody. *Dev Biol* **101**(1): 19-27.

Yoshiura, K., Kanai, Y., Ochiai, A., Shimoyama, Y., Sugimura, T. and Hirohashi, S. (1995). Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. *Proc Natl Acad Sci U S A* **92**(16): 7416-9.

You, L., Kim, J., He, B., Xu, Z., McCormick, F. and Jablons, D. M. (2006). Wnt-1 signal as a potential cancer therapeutic target. *Drug News Perspect* **19**(1): 27-31.

Young, P., Deveraux, Q., Beal, R. E., Pickart, C. M. and Rechsteiner, M. (1998). Characterization of two polyubiquitin binding sites in the 26 S protease subunit 5a. *J Biol Chem* **273**(10): 5461-7.

Yu, H., Mashtalir, N., Daou, S., Hammond-Martel, I., Ross, J., Sui, G, Hart, G. W., III, F. J. R., Drobetsky, E., Milot, E., Shi, Y. and Affar, E. B. (2010). The ubiquitin carboxyl hydrolase BAP1 forms a ternary complex with YY1 and HCF-1 and is a critical regulator of gene expression. *Mol Cell Biol* **30**(21): 5071-85.

Zeng, X., Tamai, K., Doble, B., Li, S., Huang, H., Habas, R., Okamura, H., Woodgett, J. and He, X. (2005). A dual-kinase mechanism for Wnt co-receptor phosphorylation and activation. *Nature* **438**(7069): 873-7.

Zhang, D., Chen, T., Ziv, I., Rosenzweig, R., Matiuhin, Y., Bronner, V., Glickman, M. H. and Fushman, D. (2009). Together, Rpn10 and Dsk2 can serve as a polyubiquitin chain-length sensor. *Mol Cell* **36**(6): 1018-33.

Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E. and Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. *Cell* **118**(1): 57-68.

Zhang, L., Zhou, F., Drabsch, Y., Gao, R., Snaar-Jagalska B. E., Mickanin, C., Huang, H., Sheppard, K., Porter, J. A., Lu, C. X., Dijke, P. T. (2012). USP4 is regulated by AKT phosphorylation and directly deubiquitylates TGF-beta type I receptor. *Nat Cell Biol* **14**(7): 717-26.

Zhao, B., Schlesiger, C., Masucci, M. G. and Lidnsten, K. (2009). "The ubiquitin specific protease 4 (USP4) is a new player in the Wnt signalling pathway." *J Cell Mol Med* **13**(8B): 1886-95.

Zheng, N., Schulman, B. A., Song, L., Miller, J. J., Jeffrey, P. D., Wang, P., Chu, C., Koepp, D. M., Elledge, S. J., Pagano, M., Conaway, R. C., Conaway, J. W., Harper, J. W. and Pavletich, N. P. (2002). Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature* **416**(6882): 703-9.

Zheng, N., Wang, P., Jeffrey, P. D. and Pavletich, N. P. (2000). Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell* **102**(4): 533-9.

Zhu, P., Zhou, W., Wang, J., Puc, J., Ohgi, K. A., Erdjument-Bromage, H., Tempst, P., Glass, C. K. and Rosenfeld, M. G. (2007). A histone H2A deubiquitinase complex coordinating histone acetylation and H1 dissociation in transcriptional regulation. *Mol Cell* **27**(4): 609-21.

Zhu, X., Menard, R. and Sulea, T. (2007) High incidence of ubiquitin-like domains in human ubiquitin-specific proteases. *Proteins* **69**: 1-7.

Appendice

Appendix 3.1.

Appendix 3.1 MCF7 cell aggregation at 2 hour time point.

Appendix 3.1 MCF7 cells aggregation at 2 hour time point. MCF7 cells were transfected with non-targeting siRNA or siRNA against E-cadherin or USP38 at 40nM. After 72 hours, cells were detached from wells by incubation with 2ml of 2mM EDTA/PBS for 2 mins at 37˚C. Cells were further disaggregated by pipetting up and down,and cell number was determined using a haematocytometer. Cells were then resuspended in full DMEM at 5 x 10^4 cells/ml. 20µl of the cell suspension was pipetted on the inner side of the lid of a petri dish and the lid was inverted onto a petri dish containing sterile water. The whole set up was incubated at 37˚C. Images of cells were acquired at 2 hour time point.

Appendix 5.1. Primers to amplify esiRNA region

Table S5.1 Primers to amplify esiRNA region.

DUBs	Reference sequence	Primers (5' – 3')**		Efficiency / %	Amplicon size
CYLD	NM_015247	F	TGATGAAGATTGTGGCGTGTT	55.2	418bp
		R	TGAACCTTTGTCCCCAACACC		
DUB3*	NM_201402	F	CTGAAGACACAGACAGGCGAG	45.1	406bp
		R	CACTGGCACACAAGCAGAGC		
DUB4*	NM_212553	F	CCATAGAGGCAAGCAGGAAG	55.1	512bp
		R	CCAGCACAGCATAGAGGACA		
USP1	NM_003368.4	F	TTTGCTGCTAGTGGTTTGGAG	66.2	360bp
		R	TGGCTGTGTATTTCCACCAAC		
USP2A	NM_004205.4	F	CCAATGATGTGGTGAGCCCAT	64.3	449bp
		R	GCAGCGACAGCATGTTGGCT		
USP3	NM_006537.2	F	CAATGCGCTACCTTTTGGAC	49.9	518bp
		R	TGTCTAGGCCTCTCAGTGGAA		
USP4	NM_003363.3	F	TCCACCTCAAACGTTTCTCC	56.5	358bp
		R	CCTCCATCAGAGGAACCAGA		
USP5	NM_003481.2	F	ATGCCCAGGAGTTCTTCCTT	43.1	445bp
		R	ACCCAGTCTAAGCCGAAGGT		
USP6	NM_004505.1	F	CCGTTGGAATCAACAGCAGC	51.1	498bp
		R	CCTCCTCAGGCAGATAAAGG		
USP7	NM_003470.2	F	TACGTGACTTGCTCCCAGTT	58.0	413bp
		R	ACCTGGGCCATCCCTATAAC		
USP8	NM_005154	F	TCTGGACCAGCTCTTACTGGA	62.9	427bp
		R	GTTTCTGCCAGGCATGTTCT		
USP9X	NM_004652, NM_021906	F	GAGCAGTACTAAATCTCTTGAG	66.8	351bp
		R	CATTCTGCTGAATTGGCATT		
USP9Y*	NM_004654	F	GGGCTCAGAGGTGAAACTGA	56.1	526bp
		R	AAGCTCGCTAATTCCCTCCT		
USP10	NM_005153.2	F	GCTGATCAATAAAGGGAACTGG	53.3	373bp
		R	CAACATTTCCTCATGAAGTCCA		
USP11	NM_004651.3	F	AAGCTGGTTCCTTGTGGAGA	47.3	489bp
		R	TCGAGAACCGTGATGTGTGT		
USP12	NM_182488.2	F	GTCTCCAAATTCGCCTCCATC	56.1	492bp
		R	CGTTGGGTCTGGTGTGCTG		
USP13	NM_003940.2	F	AGATCGCCTGATGAACCAAT	51.3	439bp
		R	TGCATTGGCATTATTCTCCA		
USP14	NM_005151.3	F	TGGAGTTACCATGTGGATTGA	57.4	351bp
		R	TTGCTTCCAATTTCTGTTGC		
USP15	NM_006313.1	F	TCCGCTTACCAAACCTATGC	59.7	551bp

		R	CCTGGCTGGATTCATCATCT		
USP16	NM_006447.1	F	ACCTAACGGAAGGGAGCAAT		
		R	TGTAAACATGCTTCCTTTCACC	62.5	351bp
USP17*	NM_001105662.1	F	CTCCAAGACGTTAACTTTACAC		
		R	GGTGGTCTCTCTTGAGCTCT	58.0	449bp
USP18	NM_017414.3	F	GTGCTTCTGTGAGAACTGTG		
		R	GCACTCCATCTTCATGTAAACC	59.1	436bp
USP19	NM_006677.1	F	CCCACAGTGCAAACAGCAC		
		R	GGCTCTCGTCTACCGTTGTC	42.0	356bp
USP20	NM_006676.6	F	CAAGCTCCATTCAGCCATCT		
		R	TGGCAGTAGGCGATGTAGTG	32.0	489bp
USP21	NM_012475.4	F	CGGTCCTCTGAGCCTTTCTA		
		R	TTCAGGAACTCTTGGGCATC	44.4	356bp
USP22*	NM_015276.1	F	GCTCACTATGAAGAAACTGCC		
		R	CTACTCGTATTCCAGGAACTG	52.1	394bp
USP24*	NM_015306	F	CACAGAATTATCTTTTTAATGCTTG		
		R	GGGAGCTGGTATAAATGAGTGAA	63.4	353bp
USP25	NM_013396.3	F	TTCAGGAAAATCAGGCCAAG		
		R	GCAGCAACTGATGAAGGTGA	60.8	474bp
USP26	NM_031907.1	F	TGAGCACAAGACTTCCGTTG		
		R	CCAGCTTTGTGGAACTGACA	55.2	522bp
USP27X*	NM_01145073	F	GGAAGCAGTGCCAAAATCAAATG		
		R	TCAGTAGGCTTGTGTGTTCAT	66.6	489bp
USP28	NM_020886.2	F	GCTCCACGAACAGTCACAGA		
		R	AAACCGCCAGTTATCCTCCT	64.4	480bp
USP29	NM_020903.2	F	TGTAAGCAGAAGAGTTGTGTTGC		
		R	CTTTGTAGGTCGGCATTCTTG	73.6	350bp
USP30	NM_032663.3	F	TGAAGAACAGGATGCTCACG		
		R	TCCCTTGGCTTCAATCTTTG	52.0	412bp
USP31*	NM_020718.3	F	TCTGAAGCCACCATCAGAGA		
		R	CAAAAATGCAGTCGCTTTCA	47.1	419bp
USP32	NM_032582.3	F	GAATGCTCCACGTGGTGGATG		
		R	CCTGCAACAGGCTGATAACC	71.0	434bp
USP33	NM_015017, NM_201626, NM_201624	F	AAATTCTGAAGGCGAATTTGA		
		R	CAGACACCCTGTCACAAGTCA	65.2	365bp
USP34*	NM_014709	F	GGATCCCGTACCACTTAGAC		
		R	CCTCATTGCTAGATCCATCAC	70.5	470bp
USP35*	NM_020798.2	F	CGGATCTGTATGAGCCTGTCAT		
		R	CGAGCTGGACTGCTTGAGTTTC	43.1	491bp

USP36	NM_02590.3	F	GAAGATCACCAAGGATGTAGG	57.1	475bp
		R	GGTGGAGCCATTCCTGGATAT		
USP37*	NM_020935	F	GGTAATCCGGGTAGAGGATC	72.8	442bp
		R	GAGGAACTGGCTGAGGAAG		
USP38	NM_032557.5	F	GAGGCGTTCCATTTGATTGT	58.1	452bp
		R	GCCAGAAAGGCAAAAAGATG		
USP39	NM_006590.2	F	CAACCTCCACACCCTCAAGT	58.8	406bp
		R	AGCATCTCATGGGGAGACAC		
USP40*	NM_018218	F	TAAACCCGATGCAAAGGTTC	60.6	429bp
		R	ACCAGCCTGTCACAAGTTCC		
USP41	XM_036729	F	CAGTGTCAGTACGTGCATCC	53.4	519bp
		R	GGTGAGCATGCTGCTGTTTC		
USP42	NM_032172.2	F	TTTATCGGACCACAGCTTCC	61.7	365bp
		R	ATCGTAGATGCGTTCGAGGT		
USP43*	XM_934221.1	F	ACCTTACCGGAGTCCAGCTT	50.5	355bp
		R	ACCAGCCTCATCTTGTTTGC		
USP44	NM_032147.2	F	TTCTGGGACTTGTCATTGGA	55.5	438bp
		R	GTAGTGCCCTGAGCCAAATC		
USP45	NM_001080481.1	F	GCAATTGCTCATTTCTGCTG	68.3	419bp
		R	CATAGAAAAGAAGGTAGGCTTG		
USP46	NM_022832.3	F	CAGGGAACGCTTACCAATGA	61.1	536bp
		R	GATATATCTGACGTCAGGCC		
USP47	NM_017944	F	TAGGTGGCGTCAAGTCAACA	63.5	401bp
		R	CCGATCCAGGAGTTTCCATA		
USP48	NM_001032730.1	F	TCAGCGCAGTCCTCATACAC	62.1	352bp
		R	GCCATTTCAATACACCACTCC		
USP49	NM_018561.3	F	AGCTGCTCAGTCAGGTCACAT	64.0	536bp
		R	ACCTCCCTCTGTGTTGTAGC		
USP50	NM_203494.3	F	TCACTGGCTTGTGGAACTTG	54.4	432bp
		R	CAAACAGCTGGGTGATGATG		
USP51*	NM_201286	F	CTACCAGCGTTTCGTTTGGAG	63.5	513bp
		R	GGAATATGGGTAAGTGCCTGG		
USP52	NM_014871.3	F	GATTACCTGCTGGATGAGAATG	55.5	503bp
		R	CCTGTGGGTTCACAGAATTGG		
USP53	NM_019050.2	F	TCTAGCAAGGATCCGAGTTTT	56.9	359bp
		R	CCACGTTCTTTATTTAGGTTGC		
USP54	NM_152586.3	F	GGTCGTGGTAGTGTACAAGG	62.4	547bp
		R	GATTGCAAAGGGAAGTGGTGG		
USPL1	NM_005800.4	F	CACATGCTCATGCTGCTTCAG	68.4	546bp
		R	GGTCCTGTCTCAGATGTGTG		
BAP1	NM_004656.2	F	CTCGTGGAAGATTTCGGTGT	43.4	516bp
		R	GTAGACCTTCAGCCCATCCA		
UCHL1	NM_004181.3	F	GCAGATTGAAGAGCTGAAGG	62.0	451bp

		R	AGAGAAGCGGACTTCTCCTT		
UCHL3	NM_006002.3	F	CGAGGTCACCAACCAGTTTC		
		R	GTCTGACCTTCATGGGCACT	62.1	432bp
UCHL5	NM_015984.2	F	GCTAAGCAGGTAATTAATAATGC		
		R	GGCAACTTCTGACTGAATAGC	71.4	555bp
ATXN3*	NM_004993.5	F	GGGCATTTTTCTTGGATCTTT		
		R	CAGAGTTTAGGAACGCACCA	67.7	405bp
ATXN3L*	NM_001135995.1	F	CCCATGGTGTCCATTCGGTT		
		R	TCTGTAGGTATGCCGGAAGATG	65.7	440bp
JOSD1	NM_014876.5	F	CCATGGTGACACCTCACAAG		
		R	ACACATCGGTCCTCCAACTC	46.1	408bp
JOSD2	NM_138334.2	F	CTGGAGCTGTGTGCTGTCC		
		R	TGTAGTAGACACCGTCCACCTG	10.2	352bp
BRCC3	NM_024332.3	F	AGGTGGTTCATCTCGAGTCTG		
		R	TCCAGAAGTCTCGTGGACCA	63.1	549bp
CSN5	NM_006837.2	F	TGCTCAATCAGCAGTTCCAG		
		R	ATTTTCGGTCATGCGTTTCT	62.8	420bp
CSN6	NM_006833	F	ACTCCTTTGAGCTGCTGTCC		
		R	CAGAGGCCTTGACGTACTCC	47.3	483bp
EIF3H	NM_003756.2	F	CCGTGAAGCAAGTGCAGATA		
		R	ATGCTTCAGGGGAAAAATCC	61.0	456bp
EIF3S5	NM_003754	F	CATCAAAGCCTACGTCAGCA		
		R	TTGCTGTTGAGCATGGTCTC	50.0	354bp
MPND	NM_032868.4	F	CCACACCCTGGTGGAAGTAA		
		R	TAGTCCATCTGTGCGTCGAT	25.1	351bp
MYSM1	NM_001085487.1	F	TGCATGCTCATGTTTCTATGG		
		R	CCAGGCAGGTAATCTGAGAA	69.7	363bp
PRPF8*	NM_006445	F	AACAATGTGAACGTGGCATC		
		R	AAGCACATTCTTGGGAAGGA	52.3	351bp
PSMD14	NM_005805.4	F	GCTTTGGTTGTTGGCTTTCTGG		
		R	GGTCCTGCTTGCCAACATTC	70.2	489bp
PSMD7	NM_002811.3	F	GAAAGTCAATGCCAGGGAAA		
		R	GGAGTTCAGTCCCTTCAAACC	54.8	361bp
STAMBP	NM_006463.3	F	TACTTCCGCTCTGGAGTTG		
		R	TAGGCCAGGGTCTACCTTC	65.0	471bp
STAMBPL1	NM_020799.2	F	CTGCTTTTCCACACACCAGA		
		R	GTGTGAAGATCAACGCTGGA	58.0	453bp
OTUB1	NM_017670.2	F	ATTGCTGTGCAGAACCCTCT		
		R	TCGATGAAGTGCTCGAAGAA	39.9	464bp
OTUB2	NM_023112.3	F	TGACATTCTATCCATTCTTCGG		
		R	CCATGGGCTCTACTTCGTG	45.6	476bp
OTUD1	XM_166659.7	F	GTGGAGAAGCAGGACAAGTA	45.5	527bp

*		F	GTGGAGAAGCAGGACAAGTA		
		R	TTAGACAAGGATATGGCCATAG	45.5	527bp
OTUD3	NM_015207.1	F	GGACGGCAATTGCTTGTTCAG		
		R	CAGCTTCCAGGTTCTGGACT	71.5	557bp
OTUD4	NM_199324.2	F	CAGGAATGGGTAGGACAAGTG		
		R	CTATAGACTGCAGGATTGAGTG	56.4	479bp
OTUD5	NM_017602	F	CTGTCTCTTCCGGGCTGTAG		
		R	CCATGACTCCTCCGATGTTT	55.0	460bp
OTUD6A	NM_207320.1	F	ATGGATGATCCGAAGAGTGAAC		
		R	GCAGGAACTCGTCGACGTG	28.4	559bp
OTUD6B	NM_016023.3	F	CAGGAGCCAGACATATGGAAA		
		R	CCAACCGTGTAACCGAATTA	61.0	471bp
OTUD7A*	NM_130901.1	F	GTACGAGAGCCTGGAAGAG		
		R	GGATCCACGTCACGTTCATG	41.1	434bp
OTUD7B	NM_020205.2	F	GTATAAGCTGCTGCCCTTGC		
		R	AAGCTGCCCAGTTTGTTAGC	39.9	378bp
PARP11*	NM_020367	F	GCAGTGAATTTGTGGAAGCA		
		R	AAGAGCAAACCTTCCTGCAA	59.4	461bp
TNFAIP3	NM_006290.2	F	CCCAGGAATGCTACAGATACC		
		R	GGCTTGGAGTTCAGCTTTGG	58.6	556bp
VCPIP1	NM_025054.4	F	CATGGACAGCTGAGGACTGA		
		R	TCCCAAATTCACCTGGAAAG	64.1	424bp
YOD1	NM_018566.3	F	GGATCTCAGCAATGGGGATAC		
		R	GGAGGTGTATCTGGATCAGG	66.4	552bp
ZRANB1	NM_017580.2	F	AAAACTGGCCATCTGCAATC		
		R	GTGCAAACAGAGCAAGTCCA	60.2	430bp

Table S5.1 Sequence of forward and reverse primers at the boundaries of esiRNA region.

* esiRNA region amplified from cDNA.
** All primers are flanked by the T7 promoter sequence GCTAATACGACTCACTATAGGGAGAG on their 5' ends.
Note: F – Forward primer; R – Reverse primer